AN ONTOLOGY OF MY SPIRIT, MY ALGORITHM, CATEGORY THEORY AND

Geometric Algebra

Jack Peter Kettley

Fernando Pessoa -

> “Sometimes I invent a metaphysics. With all the respectful scrupulousness of attention of someone engaged in real scientific work, it reaches the point where I may really be doing just that. The important thing is not to pride myself too much on all this, for pride is prejudicial to the precise impartiality of scientific exactitude."

I’m not quite sure of the contents here yet. Nor when I will finish this. I never know how or where to begin, and I want the reader to know; that I have sat myself here to write without a clear idea of the language or events I will use to convey what it is that I want to express. And this meta-introduction, in all its immaturity speaks to what I will present—in both its immaturity and its honesty. Writing for myself is a fleeting thing, but I carry with me a very well-established and clear intention: to present an ontological and complete body of work that captures what I have learned throughout this. What I have learned and what I will depart to put forward are not answers; rather, my understanding of this, what this is. And in it, my own primal understanding nonetheless grows, and, in itself, forms a part of my current understanding.

Both this, understanding, and their isomorphisms, I do now think to, in a deeply personal sense, ideas which may extend to a form of Universal or a form of Total. Whether by that I mean personally or collectively, whether what I may find to be universal may be merely relatable to whoever reads this; or something indigestible; or grotesque to another. And to that I am content. And what this is: I mean, and I do mean, my personal understanding, in my historical, epistemological, algebraic, and dialectical sense.

I will attempt to present— what this writing would inherently lead to, if captured— Personally an ontological body of my phenomenology, a part of it; of my personal understanding; or conversely of a jumbled pretentious text. This all may likely stand incomplete, unfinished, and in both an acceptance of pretentiousness I feel a faith in spending my time towards, again; at least for these fleeting moments, time towards it all.

I hate that word things. And maybe passion too, maybe I carry the idea of being passionate, when often I find myself too tired or stoned, high or bored, without means or lazy to, but perhaps it is that I have not constructed either the language or the actions, to perform such the task or thing that I am passionate about. Defining passion too, to a level that will satisfy what I am attempting to get at, I find difficult itself.

And to who I am to write such a thing. The room I now live in is different to the one where I began writing this. I only write alone in my room. The room in which I sit is two and a half by four metres, and on the second floor of a contemporary-modern apartment: in the kind of bleak, corporate, real-estate, sense. But it’s one that’s horizontally clad with an acceptable false wood that mimics the Victorian New Zealand architecture in the surrounding suburbs. Stuck on the opposite ends of the symmetrical windows that run single file across the apartments of the second floor, the highest floor and the one in which I sit, are false French window shades. The ground level encircled by concrete stilts, to the roadside, the north-western side, a large pin-coded, metal-barred, roller gate and door. The ground floor occupied is parking spaces and an enclosed staircase on the left-hand side looking in from the street lead to 4 apartments above. The first floor of each apartment: an open space dining room and longue with a sliding door at the north-western end that leads out onto very small balcony, with wooden dividers that separate each apartment’s balcony. Opposite, to the south-east, the door, and beside to the south-south-east the kitchenette stands tucked into the corner with a shiny, metallic newly bought modern appliances. Wooden surfaces lit by yellow LED lights, the temperature of which I’m unsure. A small a staircase somewhat hidden beside the kitchenette, slightly less deep into the room, leads to the two rooms and bathroom upstairs, my window at the southeastern end faces a courtyard with a building that mirrors mine across. The sides of the courtyard: two more apartment complexes, one of wood, whose rooms are tall and thin, with large windows, and a loft space I can see from my room on the right-hand side. To the north, my left-hand side, a concrete, three-storied, brick wall painted a blueish white. The apartment was rented unfurnished. I have recently bought: an Iranian hand carved wooden one-seater with sewn cushions that depict a girl and a deer, in what is quite an Art-Noveau style. Its arms: carved wooden lions. It is quite centrally in the dining room-lounge on the north-eastern side beside a wooden dining table and the table’s accompanying two wooden chairs; A large metal coffee table that sits closest to the sliding balcony doors, itself with a glass top spray-painted black underneath; A glass desk for my room that measures 1.2m by 0.6m and sits on two metal A-framed legs in the south-south-east corner, it’s legs two metal A’s stuck together by two metal bars 0.6m across, and so two of each of those for each end of the table. With my personal computer tower, the machine I write this on, underneath the glass desk, equipped with an Intel i7 10700k, which my computer reads in this moment is running at 4.3 gigahertz, that is 430000000 cycles per second. 16GBs of random-access memory, an RTX 2070 GPU with a couple cores dedicated to matrix multiplication and 3TBs of a mix of solid state and disk space, of which 10GBs remain. An “Old time”, as the second hand furniture store put it, chair, with a curved 1080p 27” Samsung monitor on eye saver mode in front of me, a pair of KRK monitor speakers almost parallel but slight facing inward at the far corners of the table, an IBM Model M keyboard, Fernando Pessoa’s Book of Disquietude in front of me tucked under the monitor; beside a penguin collection of Allen Ginsberg’s Howl, Kaddish, and other poems. Tucked between the PC Monitor and the speakers are the books: Fyodor Dostoyevsky stories printed by Progress Publishers in Moscow, A stolen copy of the Feynman Lectures of Physics Vol: 3, Quantum Mechanics from the University of Auckland library (never to be returned), Theodor Adorno’s Minima Moralia, W B Yeats, and a photobook of myself gifted to me by my grandparents, Infinite Jest, half-way off off the Southern end of the table, the book cover of Martin Margiela’s women’s collections 1989-2009 as my mouse-pad for my RGB Hyper-X mouse, the book itself to my right side, with my weed grinder and phone above it, bong on the ground. An empty can of Coke Classic, Tom Ford Noir, an empty glass of kratom, and an empty wine glass, on the ground beside me an Uber Eats Bag, a Glow-in-the-dark Louis Vuitton styled large glass bong, a broken studded Versace belt, a New-Era cap above the speaker, and my bed on the northern side of the room. The staircase and second floor are carpeted a shallow but comfortable carpet, the dining room-lounge a fake wood, most likely a more chemically modern counterpart to linoleum. The walls, white.

There are things that I do, and in moments, I find myself loving what I do. But why I do them: it's not that I am unaware of the genetic, traumatic, and quite universally, probabilistically deterministic forces that position me into such roles. It's whether the reader wants to accept this as passion itself.

Fernando Pessoa -

> "And naturally it is to find it necessary to reduce the spirit to some sort of physical matter surrounded by a space in which it can exist. All this depends on a great refinement of our inner feelings which, taken to its limit will no doubt reveal or create in us a genuine space like the space in which physical things exist but which in fact does not itself exist as a thing."

I want to talk about structure, complexity and irreducible representations. It is to me unclear whether people’s writing about such things are isomorphic to what that I interpret those very things to mean. And so naturally as to whether I’m talking about mathematics yet, I already talk about the isometry of ideas, and here I’m talking about an equivalence of ideas, or at the very least, the relationships and mappings between them.

Poincare -

> “How much a well-chosen word can economize thought.
>
> When our language is well chosen, it is astonishing to see how all the demonstrations made upon some known fact immediately become applicable to many new facts. Nothing has to be changed, not even the words, since the names are the same in the new cases.
>
> A word well-chosen very often causes the disappearance of exceptions to rules as announced in their former forms; it was for this purpose that the terms ‘negative quantities’, ‘imaginary quantities’, ‘infinite points’, have been invented. And let us not forget that these exceptions are pernicious, for they conceal laws. Very well then, one of those marks by which we recognize the pregnancy of a result is in that it permits a happy innovation in our language. The mere fact is oftentimes without interest; it has been noted many times but has rendered no service to science; it becomes of value only on that day when some happily advised thinker perceives a relationship, which he indicates and symbolizes by a word.
>
> The physicists also do it just the same way. They invented the term ‘energy’, a word of very great fertility, because through the elimination of exceptions it established a law; because it gave the same name to things different in substance, but alike in form.
>
> How often mathematicians of the past must have run across groups without recognizing them and how, believing these groups to be isolated things, they have found them to be in close relationship without knowing why. Today we would say that they were looking right in the face of isomorphic groups. We feel now that in a group the substance interests us but very little; it is the form alone which matters, and so, when we once know well a single group, then we know through it all the isomorphic

> groups; thanks to the words ‘groups’ and ‘isomorphism’, which sum in a few syllables this subtle law and make it at once familiar to us all, we take our step at once and in so doing economize all effort of thought.”

Algebra is the manipulation of symbols without (necessarily) regard for their meaning, especially in a way that can be formalized in logic, often cartesian logic.

We talk about objects, about numbers, the differences between them, how they are unique, the relationships that they share with each other, how equal and unequal. When we are talking about the counting numbers, or natural numbers $\mathbb{N} = \{1,2,3 \ldots\}$ we are talking about the countably infinite set of positive whole numbers, each of the elements of this set have the same quality that they are all whole, positive, numbers where $n + 1$ always exists. The natural numbers form a totally ordered set under the standard "less than" relation. This means that for any two natural numbers a and b, one and only one of the following is true: $a < b, a = b,\ \ or\ a > b$. The natural numbers are a countably infinite totally ordered set.

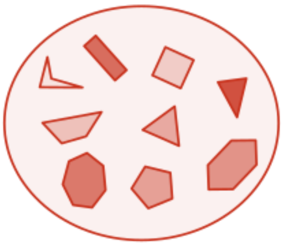
George Cantor –

> An aggregate (set) we are to understand any collection into a whole M of definite and separate objects of our intuition or our thought. The objects are called the elements of M. In signs we express x is an element of A $x \in A$. An arbitrary set M made of components 1, 2 and 3 is given $M = (1,2,3)$ We denote the uniting of many aggregates, M, N, P which have no common elements into a single aggregate (M,N,P), $M = (1,2,3), N = (4,5,6), P = (7,8,9),\ \ Q = (M, N, P) = M\ \cup N \cup P = (1,2,3,4,5,6,7,8,9)$. The elements of this aggregate are then the elements of M, of N, of P taken together.

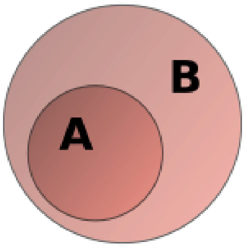
Bertrand Russel

> When mathematicians deal with what they call a manifold, aggregate, Menge, ensemble, or some equivalent name, it is common, especially where the number of terms involved is finite, to regard the object in question (which is in fact a class) as defined by the enumeration of its terms, and as consisting possibly of a single term, which in that case is the class.

A Euler diagram of a set,

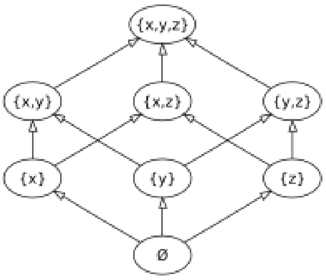


A is a subset of B, B is a superset of A. $A \subset B$



The collection of all algebraic structures of a given type will usually be a proper class. Examples include the class of all groups, the class of all vector spaces, and many others. The paradoxes of naive set theory can be explained in terms of the inconsistent tacit assumption that "all classes are sets"., one could, for example, define a class of all classes that do not contain themselves, which would lead to a Russell paradox for classes.

A partial order on a set is an arrangement such that certain pairs, for certain pairs of elements, one precedes the other, the word partial is used to indicate that not every pair of elements needs to be comparable.



The elements of the power set of {x, y, z} ordered with respect to inclusion.

Partial orders thus generalize total orders, in which every pair is comparable. A total order is some binary relation $\leq$ on some set X, which satisfies the following for all a, b and c in X.

$$a \leq a \; (reflexive)$$
$$If\ a \leq b\ and\ b \leq c\ then\ a \leq c\ (transistive)$$
$$If\ a \leq b\ and\ b \leq a\ then\ a = b\ (antisymmetric)$$
$$If\ a \leq b\ or\ b \leq a\ (strongly\ connected)$$

A well order on a set S is a total ordering on S with the property that every non-empty subset of S has a least element in this ordering. Let $(P, \leq)$ be a preordered set and let S be a subset of P. That is $S \subseteq P$. An element $l \in P$ is said to be the least element of S if $l \leq s$ for all $s \in S$. Every well-ordered set is uniquely order isomorphic to a unique ordinal number, called the order type of the well-ordered set. The well-ordering theorem, which is equivalent to the axiom of choice, states that every set can be well ordered. If a set is well ordered (or even if it merely admits a well-founded relation), the proof technique of transfinite induction can be used to prove that a given statement is true for all elements of the set.

The integers $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ can be formally constructed as the equivalence classes of ordered pairs of natural numbers $(a, b), \in \mathbb{N}$, with the equivalence relation $\sim$ with the following rule

$$(a, b) \sim (c, d) \text{ when } a + d = b + c$$

Addition and multiplication on integers can be defined in terms of equivalent operations on the natural numbers by using $[(a, b)]$ to denote the equivalence class having (a, b) as a member one has:

$$[(a, b)] + [(c, d)] \coloneqq [(a + c, b + d)]$$
$$[(a, b)] \cdot [(c, d)] \coloneqq [(ac + bd, ad + bc)]$$

The rational numbers $\mathbb{Q} = \left\{\frac{a}{b}, \middle| a, b \in \mathbb{Z}\right\}$, e.g. $\frac{1}{2}, \frac{3}{7}$ may be built as the equivalence class of ordered pairs of integers. More precisely let $(\mathbb{Z} \times (\mathbb{Z} \backslash \{0\}))$ be the set (m, n) of integers such $n \neq 0$. An equivalence relation is defined on the set by,

$$(m_1, n_1) \sim (m_2, n_2) \Leftrightarrow m_1 n_2 \equiv m_2 n_1$$

Addition and multiplication can be defined by the following rules:

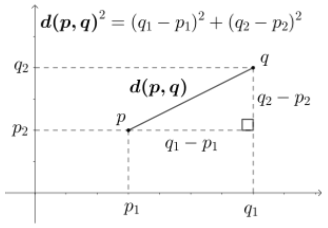
$$(m_1, n_1) + (m_2, n_2) \equiv (m_1 n_2 + n_1 m_2, n_1 n_2)$$
$$(m_1, n_1) \cdot (m_2, n_2) \equiv (m_1 m_2, n_1 n_2)$$

A metric space is a set together with a notion of distance between its elements. To see the utility of different notions of to see the utility of different notions of distance, consider the surface of the Earth as a set of points. We can measure the distance between two such points by the length of the shortest path along the surface, "as the crow flies"; this is particularly useful for shipping and aviation. We can also measure the straight-line distance between two points through the Earth's interior; this notion is, for example, natural in seismology, since it roughly corresponds to the length of time it takes for seismic waves to travel between those two points.

The notion of distance encoded by the metric space axioms has relatively few requirements. This generality gives metric spaces a lot of flexibility. At the same time, the notion is strong enough to encode many intuitive facts about what distance means. This means that general results about metric spaces can be applied in many different contexts.

Like many fundamental mathematical concepts, the metric on a metric space can be interpreted in many different ways. A particular metric may not be best thought of as measuring physical distance, but, instead, as the cost of changing from one state to another or the degree of difference between two objects.

Formally a metric space is an ordered pair (M, d), where M is a set and d is a metric on M, i.e. a function. Satisfying the following axioms for all points $x, y, z \in M$. The distance from a point to itself is zero, $d(x, x) = 0$. The distance between two distinct points is always positive: $If\ x \neq y, then\ d(x, y) > 0$. The distance from x to y is always the same as the distance from y to x $d(x, y) = d(y, x)$. The triangle inequality holds such that, $d(x, z) \leq d(x, y) + d(y, z)$, this is a natural property of both physical and metaphorical notions of distance: you can arrive at z from x by taking a detour through y but this will not make your journey shorter than the direct path.



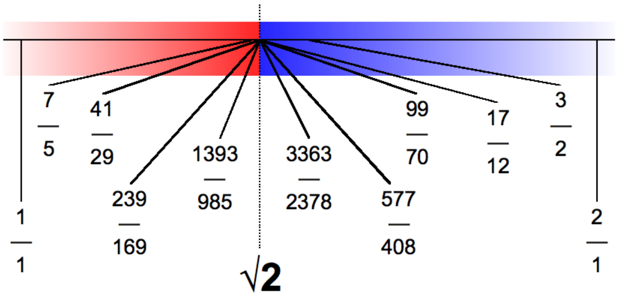
A metric space M is called complete (or a Cauchy space) if every Cauchy sequence of points in M has a limit also in M. Intuitively, a space is complete if there are no "points missing" from it (inside or at the boundary). For instance, the set of rational numbers is not complete because e.g. $\sqrt{2}$ is "missing" from it, even though one can construct a Cauchy sequence of rational numbers that converges to it. There are sequences of rational numbers that converge at the limit to irrational numbers; these are Cauchy sequences having no limit in $\mathbb{Q}$.

e.g. the sequence defined

$$x_0 = 1$$
$$x_{n+1} = \frac{x_n + 2/x_n}{2}$$

Consists of rational numbers $(1, \frac{3}{2}, \frac{17}{12}, \ldots)$; however, it converges to the irrational $\sqrt{2}$ which does not exist in $\mathbb{Q}$.

It is always possible to "fill all the holes", leading to the completion of a given space.



A Cauchy sequence is a sequence whose elements become arbitrarily close to each other as the sequence progresses.

We will consider the set of all Cauchy sequences that are defined as follows:

$$C \coloneqq \{(x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{Q}\ and\ (x_n)_{n\in\mathbb{N}}\ is\ a\ Cauchy\ sequence\}$$

For two elements, that is, for two sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$, define:

$$(a_n)_{n\in\mathbb{N}} \sim (b_n)_{n\in\mathbb{N}} :\Leftrightarrow \; (a_n - b_n)_{n\in\mathbb{N}}$$

It is a convergent sequence with limit 0, ~ is an equivalence relation on C (reflexive, symmetric, transitive), an equivalence class exists such that

$$[(x_n)_{n\in\mathbb{N}}]_{\sim} := \{[(a_n)_{n\in\mathbb{N}}]_{\sim} \mid (x_n)_{n\in\mathbb{N}} \in C\}$$

We can define the real numbers

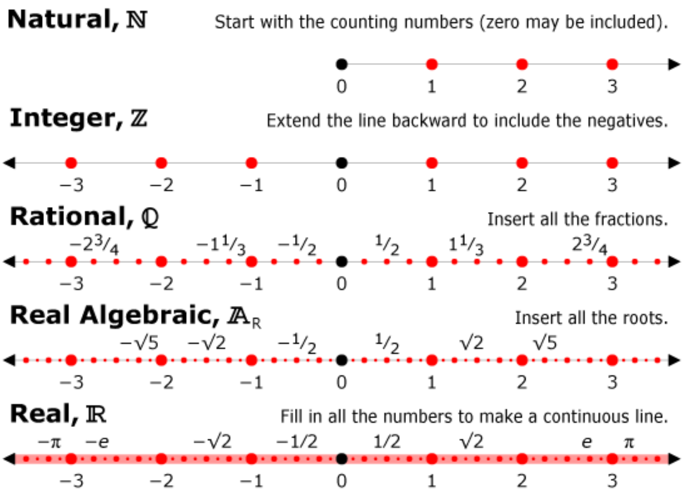
$$\mathbb{R} := \frac{C}{\sim} := \{[(x_n)_{n\in\mathbb{N}}]_{\sim} \mid (x_n)_{n\in N} \in C\}$$

This is the set of all points on the number line.

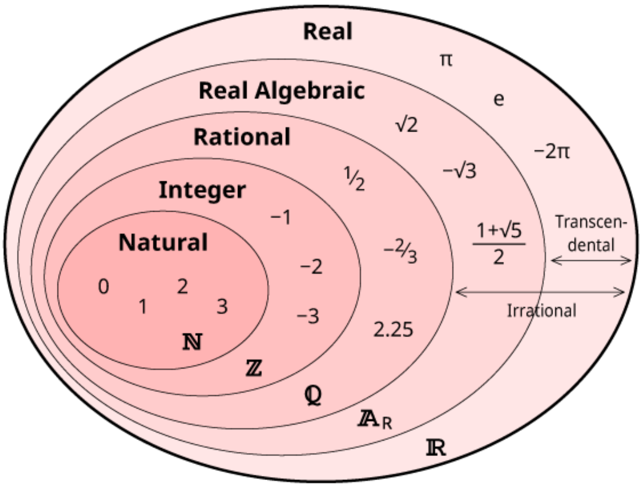
$$[(a_n)_{n\in\mathbb{N}}]_{\sim} + [(b_n)_{n\in\mathbb{N}}]_{\sim} :\Leftrightarrow \; [(a_n + b_n)_{n\in\mathbb{N}}]_{\sim} \; (well-defined)$$
$$[(a_n)_{n\in\mathbb{N}}]_{\sim} \cdot [(b_n)_{n\in\mathbb{N}}]_{\sim} :\Leftrightarrow \; [(a_n \cdot b_n)_{n\in\mathbb{N}}]_{\sim} \; (well-defined)$$

$$[(a_n)_{n\in\mathbb{N}}]_{\sim} < [(b_n)_{n\in\mathbb{N}}]_{\sim} \quad :\Leftrightarrow \quad \exists\delta > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N\colon\; \delta < b_n - a_n$$

A sequence $x_1, x_2, x_3, \ldots$ In a metric space (X, d) is called Cauchy if for every real number $r > 0$ there is a positive integer N such that for all positive integers $m, n > N, \qquad d(x_m, x_n) < r$. A metric space (X, d) is complete if any of the following equivalent conditions are satisfied. That every Cauchy sequence of points in X has a limit that is also in X. Every Cauchy sequence in X converges in X (that is, to some point in X). A distance metric for the real numbers is defined $d(x, y) = |y - x|$ by the absolute difference. The absolute value is the non-negative value regardless of sign.


Natural, ℕ
Start with the counting numbers (zero may be included).
0 1 2 3
Integer, ℤ
Extend the line backward to include the negatives.
−3 −2 −1 0 1 2 3
Rational, ℚ
Insert all the fractions.
−2¾ −1⅓ −½ ½ 1⅓ 2¾
−3 −2 −1 0 1 2 3
Real Algebraic, 𝔸ℝ
Insert all the roots.
−√5 −√2 −½ ½ √2 √5
−3 −2 −1 0 1 2 3
Real, ℝ
Fill in all the numbers to make a continuous line.
−π −e −√2 −1/2 1/2 √2 e π
−3 −2 −1 0 1 2 3

The cardinality of a set is a measure of the number of elements in a set, two sets A and B have the same cardinality if there exists a bijection (a.k.a., one-to-one correspondence) from A to B. Any set X with cardinality less than that of the natural numbers or $|X| < |\mathbb{N}|$ is said to be a finite set, e.g. a finite set $X = \{1,2,3\}$ with 3 elements. Any set that has the same cardinality as the set of natural numbers, or $|X| = |\mathbb{N}| = \aleph_0$, is said to be countably infinite. Any set with cardinality greater than that of the natural numbers or $|X| > |\mathbb{N}|$ is said to be uncountable. The real numbers $\mathbb{R}$ are uncountable, any continuous range in the real numbers is also infinite. e.g. the subset of real numbers $X = \{\, x \mid x \in \mathbb{R},\; a \leq x \leq b\}$ generated from the range (a, b) is infinite, whereas any subset of the natural numbers is finite.



Fernando Pessoa -

> "The gloriously infinitesimal remains where it is, never goes beyond what it is, and it lives free and independent. The useless and futile create intervals of humble static in our real lives. The mere insignificant existence of a pin stuck in a piece of ribbon provokes in my soul all manner of dreams and wondrous delights. I pity those who do not recognise the importance of such things."

The natural conclusion that Caeiro draws from the smallness of his hometown is the one that I do, reading those words, receiving from them what I feel to be both inspiration and liberation. Because his village is small you can see more of the world there, than you can in the city, and in this sense the village is larger than the city. I am equal in size to whatever I see, not hemmed in by the size that I am. I am filled with the same conviction Pessoa was filled with, of their ability to recognise the heavens into new constellations, I am equal in size to whatever I see.

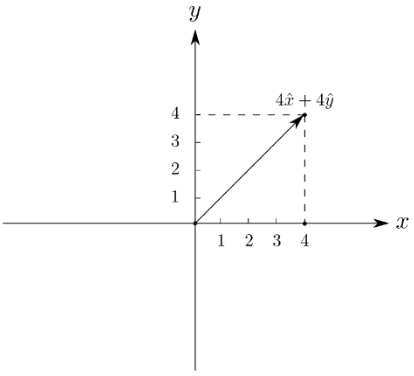
I think of my present, or the present right in front of me as a point, or an infinitesimally small point in a space, with rigour a Hilbert space. At this point, it is the summation, it is the everything that everything bears down upon me at this point, I can think of the geodesics of light through time and space piercing through this point, raging from the sun, my computer screen. I will embark latter to describe the complexity of the very structure of what I am now loosely calling a point. In this infinitesimal point of space lies the most tangled complexity, the same complexity that knits itself across everything.

Now is right where I am, and where I am, is everything now. Personally, I imagine two twin doves as I describe this, flying beside each other their whole life, eating the same, desiring the same, with the same bird song and same view each day. I point to physical separation, distance as the thing that separates our universal experience. And not just the distance between the pen and paper on my desk in a static moment in space, but the distance between the pen and paper on my desk across time.

As to my description of space as a Hilbert space rather than a Euclidean space, just space, or another common description of three-dimensional space. It is for the rigour required for this such topic, a space in which things exist has a form in itself, and to begin to describe the mechanisms of space and time it is to create abstract spaces that have properties which are useful in their ability to help with the description of the former.

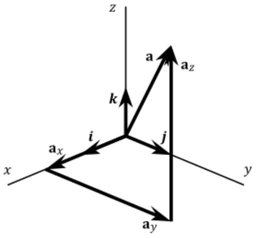
Euclidean space is intended to represent physical space. In the description of a point in three-dimensional space, its coordinate: Cartesian and Polar systems are often used. Generally, a coordinate in a Euclidean space of any dimension, that is to be able to give a description of the location of a point on a line, or a point on a sheet of paper, or a point in the general three-dimensional space around us, the Cartesian coordinates are the signed— meaning being either a positive or negative number—, distance from a central point, along perpendicular axes. For a point on a line, that would be a single number, for that on a sheet of paper, formally a plane, two coordinates (x, y), where y is perpendicular to x. and so forth for higher dimensions, (x, y, z) where z is perpendicular to y. and etc. A general way to think of any Euclidean space is a set of points satisfying certain relationships, expressible in terms of distance and angle.

Euclidean two-dimensional vector space V is equivalent to $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. The standard basis vectors of $\mathbb{R}^2$ are $\hat{x} = v_1 = (1,0)$ and $\hat{y} = v_2 = (0,1)$. Any other number in the vector space can be reached as some linear combination of the basis vectors. Any other pair of linear vectors in $\mathbb{R}^2$ also forms a basis in $\mathbb{R}^2$



The Euclidean plane $\mathbb{R}^2$ can be equipped with many different metrics, The Euclidean distance ,or Euclidean metric is defined $d_2\big((x_1, y_1), (x_2, y_2)\big) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

A point in 3-dimensional (Euclidean) vector space given in cartesian coordinates by $v = (\hat{x}, \hat{y}, \hat{z}) \in \mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ $v = a\hat{\imath} + b\hat{\jmath} + c\hat{k}$ with standard basis elements $\hat{x} = (1,0,0)$, $\hat{y} = (0,1,0)$, $\hat{z} = (0,1,0)$, these basis elements are sometimes denoted $\hat{\imath}, \hat{\jmath}, \hat{k}$ or e_0, e_1, e_2.



The set of real numbers also carries naturally the structure of a topological space and as such $\mathbb{R}$ is called the real number line. The notion of topological space aims to axiomatize the idea of a space as a collection of points that hang together ("cohere") in a continuous way. Roughly speaking, a topology on a set "of points" prescribes which subsets are to be considered "neighbourhoods" of the points they contain. Various conditions or axioms must be satisfied in order for such neighbourhood systems to form a topology, but one of the most important is that for any two neighbourhoods of a point, their intersection must also be a neighbourhood of that point. Together with the cartesian product × - the cartesian spaces $\mathbb{R}^n$, for natural numbers $n \in \mathbb{N}$, form the classical spaces of physics in n dimensions.

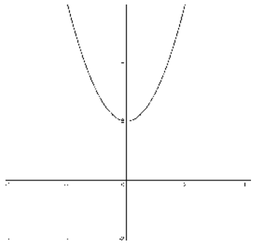
A manifold is a space that looks locally like a cartesian space, commonly a finite dimensional cartesian space $\mathbb{R}^n$, in which case one speaks of a manifold of dimension n or an n-fold, but possibly an infinite dimensional topological vector space, in which case is an infinite dimension manifold. A differentiable manifold is a topological space that is homeomorphic to a Euclidean space of some degree of differentiability. If one speaks of arbitrary or infinite differentiability, then one speaks of smooth manifolds. In any case, the type of geometry embodied in a particular flavour of manifold is controlled by a particular groupoid or, more generally, category of transformations which preserves whatever geometric features one is interested in.

That is a topological space X is called locally Euclidean if there is a non-negative integer n such that every point in X has a neighbourhood which is homeomorphic to real n-space $\mathbb{R}^n$. The real coordinate space $\mathbb{R}^n$ is an n-manifold, Any discrete space is a 0-dimensional manifold, A circle is a compact 1-manifold. Any space where there can be two distinct point x and y such that the neighbourhoods of x and y are unique is said to be a Hausdorff space. Almost all spaces encountered in analysis are Hausdorff; most importantly, the real numbers (under the standard metric topology on real numbers) are a Hausdorff space. More generally, all metric spaces are Hausdorff.

A function from a set X to Y assigns each element of X exactly one element of Y. The set X is called the domain of the function, and the set Y is called the codomain.

$$f: X \to Y$$

$$\text{e.g. } f: \mathbb{R} \to \mathbb{R},$$
$$f(x) = x^2 + 2$$



A surjective is a function f such that, for every element y of the function's codomain, there exists at least one element x in the function's domain such that f(x) = y. In other words, for a function f : X ⟶ Y, the codomain Y is the image of the function's domain X, It is not required that x be unique; the function f may map one or more elements of X to the same element of Y.

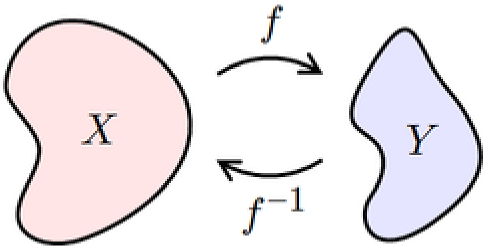
An injective function is a function f that maps distinct elements of its domain to distinct elements; that is, x1 ≠ x2 implies f(x1) ≠ f(x2). (Equivalently, f(x1) = f(x2) implies x1 = x2 in the equivalent contrapositive statement.) In other words, every element of the function's codomain is the image of at most one element of its domain. The term one-to-one function must not be confused with one-to-one correspondence that refers to bijective functions, which are functions such that each element in the codomain is an image of exactly one element in the domain.

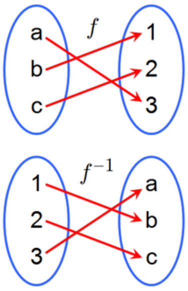
A bijective function, or one-to-one correspondence between two mathematical sets is a function such that each element of the second set (the codomain) is mapped to from exactly one element of the first set (the domain). Equivalently, a bijection is a relation between two sets such that each element of either set is paired with exactly one element of the other set.

A Homomorphism is an isomorphism if and only if it is bijective. In most cases bijective mappings are isomorphisms.

As for an isomorphism, it is the same as the inverse morphism. The multiplicative identity of X operating on X gives X. And note $I_x \cdot X = X$, for example $1 \cdot 2 = 2$. Where 1 is the multiplicative identity of the real numbers.
In all cases and using the composition of functions it is that $f^{-1} \circ f = I_x$, more simply we can say if a function operates on x and gives some number, then taking the inverse of that function and applying it on the number produced will give the original x.

$$f^{-1}(f(x)) = x, for\ every\ x \in X$$





A morphism is a map between objects. Two objects x and y are isomorphic if there exists an isomorphism from x to y (or equivalently, from y to x). A morphism that has an inverse morphism is called an isomorphism. It is an abstraction of homomorphism, a function between two algebras that preserves the structure. This itself is even an abstraction of a more descript function that has some mapping from x to y i.e. $f: x \rightarrow y$. That is that x and y share a similar structure.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g \circ f"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

The function f is a morphism between X and Y, as g is from Y to Z, the composition of functions g and f gives a morphism from X to Z

Given arbitrary functions, $f(x) = x + 3$ *and* $g(y) = y + 3$, *and an arbitrary set* $X = (1,2,3)$, *we have,* $Y = f(X) = X + 3 = (4,5,6)$, $Z = g(Y) = (4,5,6) + 3 = (7,8,9)$,

$$Z = g \circ f = g(f(x)) = ((1,2,3) + 3) + 3 = ((X + 3) + 3) = (7,8,9)$$

In category theory, the idea of functions can be abstracted further into the idea of functors, a function is an element-wise mapping from X to Y , but it does not inherently preserve or map the structure of X onto Y. A set X, barren of additional structure is comprised solely of elements $x \in X$. A function from one set to another merely sends elements from the first to the second. A category C is comprised of objects $ob(C)$ and between any two objects A, B, a class of morphisms, denoted hom (A, B). The objects A, B need not be distinct; there is always the identity morphism defined $id_A \in \text{hom}\ (A, A)$ so that $f \circ id_A = f$ and $id_A \circ g = g$ for any $f \in \text{hom}(A,\cdot)$ and any $g \in \text{hom}(\cdot, A)$.

So, we can say that if $f: X \rightarrow Y$ is a homomorphism, then the function acting on the identity of X is equal to the identity of Y, and that the function acting on the identity of Y is equal to the identity of X.

$$f(I_x) = I_y$$

Proof : take, $x \in X$,

$$f(xI_y) = f(x)f(I_y)$$
$$f(xI_y) = f(x)$$

$$([f(x)]^{-1}f(x))f(I_x) = ([f(x)]^{-1}f(x))$$
$$I_y f(I_y) = I_y$$
$$f(I_y) = I_x$$

Such that,

$$X \cong Y$$

There is a deep form of perseveration of structure, and to a degree what information can be held within that structure by the bijective mapping, the isomorphism.
In some cases,

$$X = Y, or\ X \equiv Y$$

We see that the groups are the same in a structural way, but their elements are different. Their elements may be equal. Equality is when two objects are exactly the same, and everything that is true about one object is true about the other, while an isomorphism implies everything that is true about a designated part of one object's structure is true about the other's

In cases, in lower categories, we may assume an isomorphism to be a form of equivalence or equality between two objects, groups, numbers, algebras.

And for this inverse morphism to exist, there must be underlying structure between X and Y such that they are the same. There is a need for good definitions.

Bryan Garner -

> "What's wrong with genteelism's like 'prior to' and 'subsequent to'"

David Foster Wallace -

> "Well, I have trouble parsing your question. Genteelism seems to me to be an overly charitable way to characterize them. They are to me like puff words. They are like using "utilize" instead of "use", which, in 99 cases out of 100, is just stupid. Or 'individual' for 'person'. More syllables it's just puffed up. Why say 'prior to' rather than 'before'? Everybody knows what 'before' means, it's fewer words. And I think, technically, given the Latin roots, it should be 'posterior to', if you're going to use 'prior to'. So, if you're saying 'prior to' and 'subsequent to', you are in fact, in a very high-level way, messing up grammatically. But would you ever want to say, 'posterior to'? No. But you will notice, this is the downside of starting to pay attention, you start noticing , you know, all the people who say, 'at this time', rather than, 'now'. Why did they just take up one-third of a second of my lifetime making me parse, 'at this time', rather than just saying 'now', to me. And you start getting bugged. But you get to become more careful and attentive in your own writing, so you become an agent of light and goodness, rather than the evil that's all around."

Consider the following two strings of 32 lowercase letters and digits:

$$Abababababababababababababababab, \quad and$$
$$4c1j5b2p0cv4w1x8rx2y39umgw5q85s7$$

The first string has a short English-language description, namely "write ab 16 times", which consists of 17 characters. The second one has no obvious simple description (using the same character set) other than writing down the string itself, i.e., "write 4c1j5b2p0cv4w1x8rx2y39umgw5q85s7" which has 38 characters. Hence the operation of writing the first string can be said to have "less complexity" than writing the second.

More formally, the complexity of a string is the length of the shortest possible description of the string in some fixed universal description language (the sensitivity of complexity relative to the choice of description language is discussed below). It can be shown that the Kolmogorov complexity of any string cannot be more than a few bytes larger than the length of the string itself. Strings like the "abab" example above, whose Kolmogorov complexity is small relative to the string's size, are not considered to be complex. The Kolmogorov complexity can be defined for any describable object.

An irreducible representation is a representation that has no smaller non-trivial representations "sitting inside it". Given some algebraic structure, such as a group, equipped with a notion of (linear) representation, an irreducible representation is a representation that has no nontrivial proper sub-object in the category of all representations in question and yet which is not itself trivial either. In other words, an irrep is a simple object in the category of representations. A representation that has proper nontrivial sub-representations but cannot be decomposed into a direct sum of such representations is an indecomposable representation but still reducible. In good cases for finite dimensional representations, the two notions (irreducible, indecomposable) coincide.

Be the north wind, it's the one that hits here, the most, my home, and for the poetry of it I'll say, "where the virgin laid". For it's the poetry of this narrative, and that one that can be written about the algorithm, my life. I feel, as the nights were colder before. They arrive, cats, drawn by the yoke of bulls, and before they go, they offer everything of their gentle protection. For I know it's the cradle, of what is coming. The cats always come past, the past I was shown, on the steps where played with the black cat, in a place near a place named Nelson, New Zealand. I built things a lot when I was younger—rules for games, role-playing as a child—in relation to the things that I was, and am still, passionate about: Imaginatively in my backyard I would advance from the neolithic to the contemporary. On my personal computer I would do the same, progressing from the raw earth underneath the Robinia tree, through logic and the creation of language, into thoughts themselves, advancing through feudalism, castles, to imperialism and back to my cotemporary living room. When I entered preparatory school, rockets from PVC tubes, cooking stump remover (potassium nitrate) as an oxidizer, sometimes mixed with rust power, with white sugar, sucrose, as the fuel. The cooking process over a camp burner to form a solid candy like fuel instead of powder to be stored inside the rocket. I note this cooking process led to explosion twice. And it was either the metal end of a garden hose—the same New Zealand hose that would become smaller and smaller as the years went by—and once moulded concrete… which led to explosion upon ignition. Ignition occurred by connecting a 9V battery to a switch from my local circuitry store and nichrome out of a smashed filament bulb which would ignite outside of its vacuum once connected to a power source. This hobby evolved into turbine jet engine building in high school. The PVC tubes were replaced with stainless steel, and the turbines were cut from sheets that sent the blade of my small circular saw flying, returning me to the computer. This led me to pursue a science degree and eventually embark on post-graduate study in quantum physics. My passion for linear algebra in my undergraduate studies led me to this path, though I ultimately dropped out—a decision influenced by the logic I constructed from the language in my head." And ultimately in the pursuit of my passion, which I hate to declare out of the cringe declaration of practice, the pursuit to express my feelings and perception by whatever means.

The modern concept of limit originates from Proposition X.1 of Euclid's Elements, which forms the basis of the Method of exhaustion found in Euclid and Archimedes: "Two unequal magnitudes being set out, if from the greater there is subtracted a magnitude greater than its half, and from that which is left a magnitude greater than its half, and if this process is repeated continually, then there will be left some magnitude less than the lesser magnitude set out."

The expression "0.999…" should be interpreted as the limit of the sequence 0.9, 0.99, 0.999, … and so on. This sequence can be rigorously shown to have the limit 1. And therefore, this expression is meaningfully interpreted as having the value of 1. Formally suppose $a_1, a_2, \ldots$ is a sequence of real numbers. When the limit of the sequence exists, the real number L is the limit of the sequence if an only if for every real number $\varepsilon > 0$, there exists a natural number N such that for all $n > N$ we have $|a_n - L| < \varepsilon$. The common notation

$$\lim_{n\to\infty}(a_n) = L$$

is read "the limit of a_n as n approaches infinity equals L".

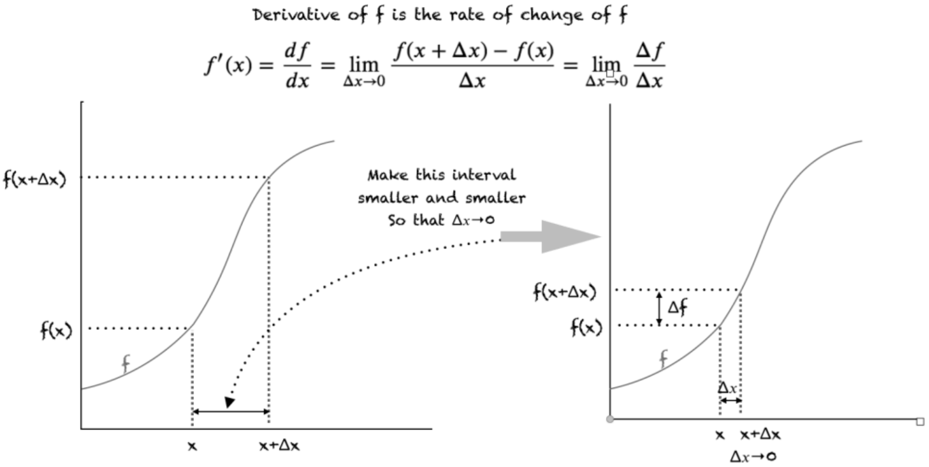
A differentiable function of one real variable function is a function whose derivative exists at each point in its domain.

A function $f : U \to \mathbb{R}$, defined on the open set $U \subset \mathbb{R}$ is said to be $differentiable$ at $a \in U$ if the derivative

$$f'(a) = \lim_{h\to 0}\left(\frac{f(a+h) - f(a)}{h}\right)$$

exists. This implies the function is continuous at a.

The derivative of a function of a single variable at a chosen input value, when it exists, is the slope of the tangent line to the graph of the function at that point. The tangent line is the best linear approximation of the function near that input value. For this reason, the derivative is often described as the instantaneous rate of change, the ratio of the instantaneous change in the dependent variable to that of the independent variable. The process of finding a derivative is called differentiation.

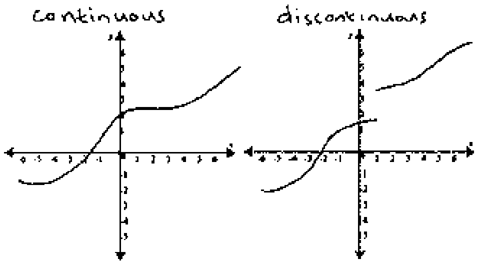

Derivative of f is the rate of change of f
f(x+Δx)
f(x)
f
x
x+Δx
Make this interval smaller and smaller So that Δx→0
Δf
Δx
Δx→0

$$f'(x) = \frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$$

Euclid -

> "The curved line a is the first species of quantity, which has only one dimension, namely length, without any width nor depth, and is nothing else than the flow or run of the point which will leave from its imaginary moving some vestige in length, exempt of any width."

This definition of a curve has been formalized in modern mathematics as: A curve is the image of an interval to a topological space by a continuous function.



Suppose a curve is given by some function $f(x)$, a curve by definition is continuous such that it has a derivative. Suppose the path of some object across time is given by $f(t)$. The derivative of f with respect to time is the velocity $f'(t) = v(t)$, the derivative of the velocity gives acceleration $f''(t) = v'(t) = a(t)$. Then that object path through time with respect to some origin is given by,

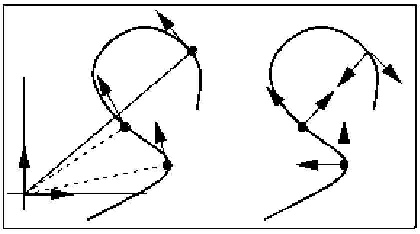
$$f(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$$

The instantaneous velocity is given by

$$v(t) = \frac{df}{dt} = \frac{dx(t)}{dt}\hat{\imath} + \frac{dy(t)}{dt}\hat{\jmath} + \frac{dz(t)}{dt}\hat{k}$$

The acceleration

$$a(t) = \frac{d^2f}{dt^2} = \frac{dv}{dt} = \frac{d^2x(t)}{dt^2}\hat{\imath} + \frac{d^2y(t)}{dt^2}\hat{\jmath} + \frac{d^2z(t)}{dt^2}\hat{k}$$

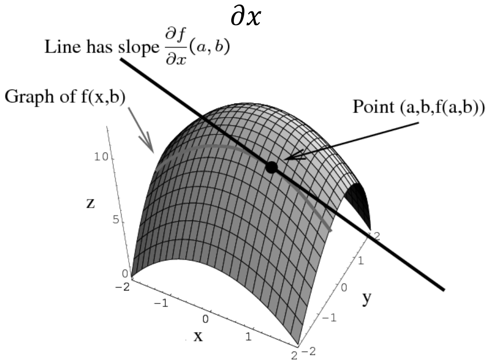


1 w.r.t some origin, 2 with intrinsic coordinates

For functions with multiple parameters say such as a function that changes independently in the x, y and z directions the derivative in respect to one of those directions is given by the partial derivative in that direction.

e.g. for a two-dimensional function, i.e. a surface. $f : x, y \to \mathbb{R}^3$, the derivative at a point (a,b) in the x direction is given by

$$\frac{\partial f(a,b)}{\partial x}$$



When we take the derivative of such a function we are taking the sum of partial derivatives,

$$\frac{dF}{dX} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}$$

The partial derivative, that is the gradient or slope of a function in some defined direction, or variables is defined as a limit. Let U be an open subset of $\mathbb{R}^n$ and $f: U \to \mathbb{R}$ a function. The partial derivative of f at the point $a = (a_1, a_2, \dots, a_n) \in U$ with respect to the ith variable x_i is defined as

$$\frac{\partial}{\partial x_i} f(a) = \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}$$

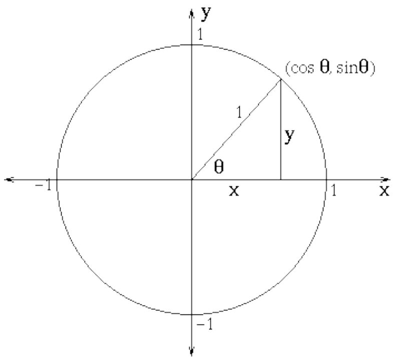
Where e_i is the unit vector of the ith variable x_i.

An important example of a function of several variables is the case of a scalar valued function $f(x_1, \dots, x_n)$ on a domain in Euclidean space $\mathbb{R}^n$ (e.g. on $\mathbb{R}^3, f(x, y, z)$) , in this case f has a partial derivative with respect to each variable (x, y, z) or (x_1, x_2, x_3) . At the point a, these partial derivatives define the vector $\nabla f(a) = (\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a))$, more generally we can abuse this notation and define, $\nabla = \sum_{j=1}^{n} [\frac{\partial}{\partial x_j}]\hat{e}_j = [\frac{\partial}{\partial x_1}]\hat{e}_1 + [\frac{\partial}{\partial x_2}]\hat{e}_2 + \dots + [\frac{\partial}{\partial x_n}]\hat{e}_n$

i. The Circle

Imagine any right-angle triangle, with its point, the one at the peak of its adjacent and hypotenuse, dead centre in a circle, and that the point at the peak formed by its opposite and hypotenuse lies on the circle.

The angle that is formed between the opposite and hypotenuse, for any of these triangles we have described, is $\sin\theta$.



We can describe any point on the circle then algebraically.

$x = r\cos(\theta)$, $y = r\sin\theta$, where x lies on some axis, and y lies on an axis, perpendicular, orthogonal, at 0 on the x axis, your standard graph.

Length, magnitude, and Euclidean norm is given by $\| \hat{x} \| = \sqrt{\hat{x}^2}$, the length of a two dimensional vector or “arrow” is given by the square root of the sum of the square of its two components, by the Pythagorean theorem $\| \hat{x} \| = \sqrt{x^2 + y^2}$ in three dimensions $\| \hat{x} \| = \sqrt{x^2 + y^2 + z^2}$ etc… . Unitary matrices and unit vectors, unit complex numbers etc. all have length / norm of 1, unit length.

|x| is the notation for the absolute operator forcing x value of x to be positive, and real. The norm is also always positive and real.

S^1, is the 1-sphere and it is the circle, the circumference of a disk, in the two-dimensional plane. Homeomorphic to the real projective line.
In two dimensional cartesian coordinates an element on the 1-sphere, i.e a point on the circumference of a circle with radius 1 can be parameterized by two coordinates, (x, y) such that $\sqrt{x^2 + y^2} = 1$

i. The Circle. A New Approach.

The imaginary unit i is defined solely by the property that its square is -1:

$$i^2 = -1$$

With i defined this way it follows that i and -1 are both square roots of -1

Given

$$i^2 = -1 \; or \; \sqrt{i} = \sqrt{-1}$$

$$\begin{gathered} i \times i = -1 \\ -1 \times i = -i \\ -i \times i = 1 \\ 1 \times i = i \end{gathered}$$

More so

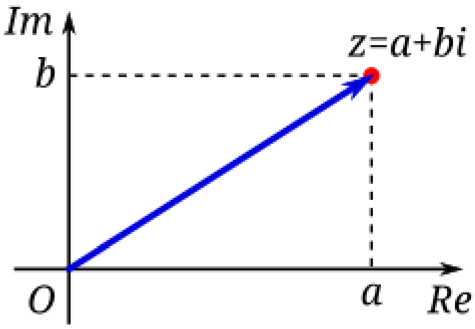
$$\begin{array}{llll} i^0 = 1, & i^1 = i, & i^2 = -1, & i^3 = -i, \\ i^4 = 1, & i^5 = i, & i^6 = -1, & i^7 = -i \\ \vdots & \vdots & \vdots & \vdots \end{array}$$

That is the powers of i are cyclic.

A complex number is an expression of the form $a + ib$

The set of all complex numbers:

$$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}, \; i^2 = -1\}$$



The complex numbers $\mathbb{C}$ can be written as ordered pairs (a, b) of real numbers a and b, with the addition operator being component wise and with multiplication defined by

$$(a,b)(c,d) = (ac - bd, ad + bc)$$

The complex number whose second component is zero is associated with a real number: the complex number $(a,0)$ is associated with the real number a.

The complex conjugate $(a,b)^*$ of (a,b) is given by $(a,b)^* = (a^*, -b) = (a, -b)$, since a is a real number and is its own conjugate. The conjugate has the property that

$$(a,b)^*(a,b) = (aa + bb, ab - ba) = (a^2 + b^2, 0)$$

Which is a non-negative real number. In this way conjugate defines a norm, making the complex numbers a normed vector space over the real numbers: the norm of a complex number z is $|z| = (z^* z)^{\frac{1}{2}}$.

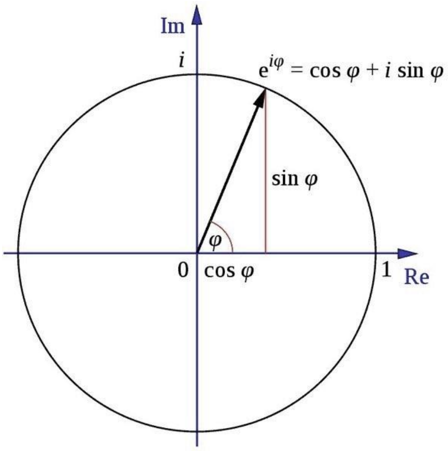
Given a cartesian coordinate system, the unit n-sphere can be defined as $S^n = \{x \in \mathbb{R}^{n+1} : \| x \| = 1\}$.where n is the dimensionality of the sphere. this reads as the n-sphere is equal to the set of all objects in $\mathbb{R}^{n+1}$ with a magnitude equal to one.

This is a very common structure in 2 and 3 dimensions, the circle and the sphere, the n-sphere generalizes the geometric structure of the sphere and extends it to any n dimensions.

, it can be also thought of as a vector with magnitude, length of 1 parametrized by an angle θ, between 0 and 360 degrees or $\phi \in (0, 2\pi)$ radians around the origin, this is isomorphic to a unit complex number, given Euler's identity $e^{i\pi} + 1 = 0$ and $e^{i\phi} = cos\phi + isin\phi$. In the complex plane, which is a special interpretation of the common cartesian x-y plane, i is the point located one unit away from the origin along the imaginary axis which is orthogonal to the real axis.

i.e. the circle group $\mathbb{S}^1$ is equivalent to the unit complex numbers $\mathbb{T}$, or the unit real numbers in $\mathbb{R}^2$.

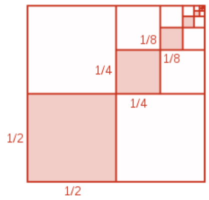
$$\mathbb{S}^1 = \{x \in \mathbb{R}^2 : \| x \| = 1\} \cong \mathbb{T} = \{z \in \mathbb{C} : \| z \| = 1\},$$



As a complex number consists of two independent real numbers, they form a two-dimensional vector space over the real numbers. Besides being of higher dimension, the complex numbers can be said to lack one algebraic property of the real numbers. Every Dedekind complete totally ordered field is isomorphic to the reals, where squares are necessarily non-negative in an ordered field. This implies that the complex numbers cannot be ordered since the square of the imaginary unit i is -1.

An example of a geometric series found by Archimedes.

$$\frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \frac{1}{4^4} \dots = \sum_{n=1}^{\infty} \frac{1}{4^n} = \frac{1}{3}$$



$$a + ar + ar^2 + ar^3 + \cdots = \frac{a}{1-r}$$

Where a = ¼ and r = ¼ , $\frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{1}{3}$

That is,

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots$$

which is valid for all $|x| < 1$

A Maclaurin series is an infinite series of the form:

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

And we see the similarity in the taylor expansion to that of sine and cosine, in fact we see their equality:

$$z = a + ib$$

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

$$\begin{aligned} e^{ix} &= 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} + \frac{(ix)^6}{6!} + \frac{(ix)^7}{7!} + \frac{(ix)^8}{8!} + \cdots \\ &= 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} - \frac{x^6}{6!} - \frac{ix^7}{7!} + \frac{x^8}{8!} + \cdots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots\right) \\ &= \cos x + i \sin x, \end{aligned}$$

The group of rotations in 2D is SO(2) = S^1, that is the orthogonal group O(n) being the group of distance preserving transformations of an n-dimensional space and the special orthogonal group SO(n) which contains the identity element. SO(2) is isomorphic to the group U(1) of complex numbers of unit length.

Multiplication on the circle group is equivalent to the addition of angles, therefore it is the multiplicative group of all complex numbers with an absolute value (consider this intuitively as length or radius) of 1, simply the unit complex numbers.

This follows immediately from the fact the mapping $e^{i\phi} \mapsto \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$ is the group isomorphism.

Geometrically observe that U(1) is the unit circle S^1 ($SO(2) \cong U(1) \cong S^1$). If you view the real and complex numbers purely as sets ignoring all possible structure on them, then the elements of $\mathbb{R}^2$ and $\mathbb{C}$ are isomorphic in the category of sets due to the bijection $(a, b) \leftrightarrow a + bi$. In real vector spaces $\mathbb{R} \times \mathbb{R}$ has the basis $[(1,0), (0,1)]$ or $\{\hat{x}, \hat{y}\}$and $\mathbb{C}$ has a (real) basis $\{1, i\}$, since these real vector spaces both have dimension 2, they are isomorphic. Every plane rotation ρ_θ by an angle θ is represented by multiplication by the complex number $e^{i\theta} \in U(1)$, in that sense for any two complex numbers $z, z' \in \mathbb{C}$
$z^{'} = \rho_\theta(z).\ iff\ \ z^{'} = e^{i\theta}z.$

Integration

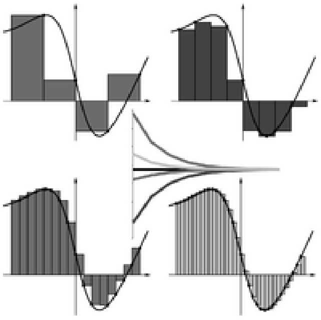
In the one dimensional case, with respect to integration in the simplest terms we want to find the area under some curve, from an interval (a through to b), sometimes we want to find the area under a curve that interval is infinite, and sometimes we want to find the area under a curve whose parts are both positive and negative, for all that means, "negative area". It is also thought of as an antiderivative, Velocity is the integral to acceleration, it's the antiderivative to acceleration, for if acceleration represents the rate at which velocity is changing at each point in time, then its velocity is the sum of its acceleration across all points in time, or between some interval.

"I have to pay a certain sum, which I have collected in my pocket. I take the bills and coins out of my pocket and give them to the creditor in the order I find them until I have reached the total sum. This is the Riemann integral. But I can proceed differently. After I have taken all the money out of my pocket, I order the bills and coins according to identical values and then I pay the several heaps one after the other to the creditor. This is my integral." (letter from Henri Lebesgue to Paul Montel)

The Riemann sum, let f be a real-valued function defined on the interval $[a,b]$ The Riemann sum of f with respect to the tagged partition $x_0,\dots,x_n$ together with $t_0,\dots,t_{n-1}$ is

$$\sum_{i=0}^{n-1} f(t_i)(x_{i+1}-x_i)$$

Each term in the sum is the product of the value of the function at a given point and the length of an interval. Consequently, each term represents the (signed) area of a rectangle with height $f(t_i)$ and width $x_i + 1 - x_i$.



Loosely speaking, the Riemann integral is the limit of the Riemann sums of a function as the partitions get finer. If the limit exists then the function is said to be integrable (or more specifically Riemann-integrable).

"This part is sometimes referred to as the first fundamental theorem of calculus." (Apostol 1967)

Let f be a real-valued function defined on the interval $[a,b]$. Let F be the function defined by

$$F(x) = \int_a^x f(t)dt$$

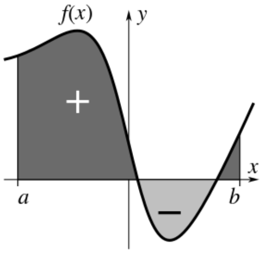
The F is uniformly continuous on [a, b] and differentiable on the open interval (a,b) and

$$F'(x) = f(x)$$

For all x in (a,b) so F is an antiderivative of f.

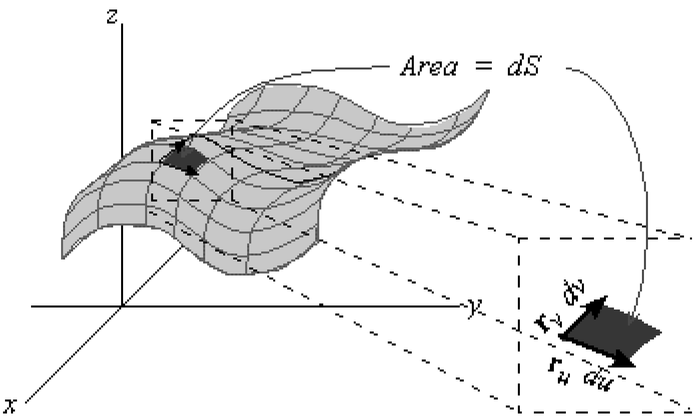
It’s Corollary,

$$\int_{\boldsymbol{a}}^{\boldsymbol{b}} f'(x)dx = F(b) - F(a)$$



The two-dimensional case we have

$$A = \iint_S dS$$



An $n \times m$ **matrix**, is a matrix with n rows and m columns A can be represented as

$$A = A_{nm} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

If A is an $m \times n$ matrix and B an $n \times p$ matrix, the **matrix product** $C = AB$ is defined to be the $m \times p$ matrix such that

$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, p$.

That is, the entry c_{ij} of the product is obtained by multiplying term-by-term the entries of the ith row of A and the jth column of B, and summing these n products. In other words c_{ij} is the dot product of the ith row of A and the jth column of B. Thus the product of AB is defined if and only if the columns of A equals the number of rows in B.

The identity matrix for $n \times n$ matrices is given by $I, AI = A, I^2 = I$, e.g. $I_{2\times2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $I_{3\times3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Given an arbitrary matrix $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$,

$$I_2 B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$BI_2 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Matrix Multiplication

$$\begin{bmatrix} \mathrm{a} & b \\ \mathrm{c} & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} ax + by + c \\ dx + ey + f \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}$$

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} ax + by + cz + d \\ ex + fy + gz + h \\ ix + jy + kz + l \\ 1 \end{bmatrix}$$

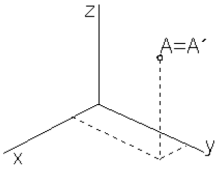
The Euclidean group is the group of (Euclidean) isometries of a Euclidean space $\mathbb{E}^n$; that is, the transformations of space that preserve the Euclidean distance between any two points, that is the group of Euclidean transformations. The group depends only on the dimension n of space and is denoted $ISO(n)$ for inhomogeneous special orthogonal group. The Euclidean group $E(n)$ comprises all translations, rotations, and reflections of $\mathbb{E}^n$; and arbitrary finite combinations of them. The Euclidean group can be seen as the symmetry group of the space itself and contains the group of symmetries of any figure (subset) of that space. The continuous trajectories in $ISO(3)$ play an important role in classical mechanics, because they describe the physically possible movements of a rigid body in three-dimensional space over time. One takes $f(0)$ to be the identity transformation I of $\mathbb{E}^3$ which describes the initial position of the body. The position and orientation of the body at any later time t will be described by the transformation $f(t)$. Since $f(0) = I$ is in $E^+(3)$, the same must be true of $f(t)$ for any later time. For that reason, the direct Euclidean isometries are also called "rigid motions".

Euclidean metric **transformations** preserve invariant size of the figure, they only change the figure position in space. Relation between coordinates of the original and its image in the given transformation is expressed by equations of the transformation. The matrix of linear transformations is the matrix of the system of equations

$$A = [x, y, z, 1]$$

$$\begin{gathered}(x, y, z, 1) \to (x', y', z', 1)\\ x' = f(x, y, z)\\ y' = g(x, y, z)\\ z' = h(z, y, z)\\ A' = A \cdot T\end{gathered}$$

The identity is given by $A' = A \cdot I_4$ $\qquad T = I_4 = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$

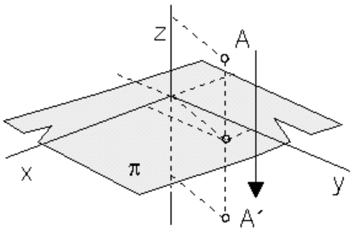


$\mathbf{A' = A\,.\,T}_I$

For example, symmetry with respect to the coordinate plane $\pi = xy$, can be given by $A' = AT_{S\pi}$

$$T_{S\pi} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

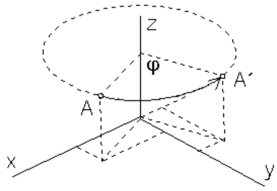
$$x' = x, \qquad y' = y, \qquad z' = -z$$



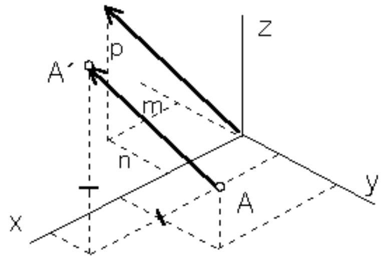
Rotation around a coordinate axis z by angle ϕ, $A' = A \cdot T_{O_z}$

$$T_{O_z} = \begin{bmatrix} \cos\phi & \sin\phi & 0 & 0 \\ -\sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$x' = x\cos\phi - y\sin\phi\,,$$
$$y' = x\sin\phi + y\cos\phi, \qquad z' = z$$



Translation by a vector v = $(\boldsymbol{m}, \boldsymbol{n}, \boldsymbol{p}, \boldsymbol{0})$



$$A' = A \cdot T_v = A \cdot \begin{bmatrix} 1 & 0 & 0 & m \\ 0 & 1 & 0 & n \\ 0 & 0 & 1 & p \\ 0 & 0 & 0 & 1 \end{bmatrix} = [x, y, z, 1] \begin{bmatrix} 1 & 0 & 0 & m \\ 0 & 1 & 0 & n \\ 0 & 0 & 1 & p \\ 0 & 0 & 0 & 1 \end{bmatrix} = [x + m, y + n, z + p, 1]$$

$$x' = x + m$$
$$y' = y + n$$
$$z' = z + p$$

The general affine transformation is given by, $|T_A| = 0$

$$T_A = \begin{bmatrix} a & b & c & m \\ d & e & f & n \\ g & h & i & p \\ 0 & 0 & 0 & s \end{bmatrix}$$

$$x' = ax + dy + gz + m$$
$$y' = bx + ey + hz + n$$
$$z' = cx + fy + iz + p$$
$$h' = s$$

$$A = (x, y, z, 1) \rightarrow \; A' = (x', y', z', h') = A \cdot T_A = s\left(\frac{x'}{s}, \frac{y'}{s}, \frac{z'}{s}, 1\right)$$
$$= s\left(\frac{ax + dy + gz + m}{s}, \frac{bx + ey + hz + n}{s}, \frac{cx + fy + iz + p}{s}, 1\right)$$

The coefficients a through to i, m n, p, s $\neq 0$ have a clear geometric interpretations:

- a, e, i are nonzero scales on coordinate axis x,y,z
- s coefficient of scaling to the origin
- m, n, p – coordinates to the translation vector
- b, c, d, f, g, h – coefficients to the general affine transformation.

Affine transformations do not preserve invariant the size of line segments and angles.
Special subsets of affine transformations form similarities, preserving invariant the size of angles.
Any Euclidean transformation is an affine transformation.

The Galilean Transformation is given by:

$$x' = x - vt$$
$$y' = y$$
$$z' = z$$
$$t' = t$$

Note that the last equation holds for all Galilean transformations up to addition of a constant and expresses the assumption of a universal time independent of the relative motion of different observers.

In the one-dimensional case we can show that newton's second law is invariant under a Galilean transformation.

$$F = m\frac{d^2x}{dt^2}$$

The Galilean transformation in one dimension, with velocity in the positive x direction given by:

$$x = x' + vt'$$
$$t = t'$$

Also note quickly rearranging this formula gives us the definition of velocity $\frac{x-x'}{t'} =$ v, that is, the change in distance over some global time that has taken place is our velocity.

Let F' be the force in the primed coordinates.

$$\frac{d^2x}{dt^2} = \frac{d^2}{dt^2}(x + vt') = \frac{d^2x'}{dt^2}$$
$$F' = m\frac{d^2x'}{dt^2} = m\frac{d^2x}{dt^2} = F$$

Note that m does depend on velocity relativistically, so the 2nd law and/or notion of invariance must be changed in relativity. In the relativistic case under the Lorentz transformation time is dependent on velocity and position and with reference to the speed of light.

Again, in the one-dimensional case with velocity in the x direction.

$$x' = \gamma(x - vt)$$
$$t' = \gamma\left(t - \frac{vx}{c}\right)$$
$$where, \qquad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

When speed v is much smaller than c, the Lorentz factor is negligibly different from 1, but as v approaches c, γ grows without bound. The value of v must be smaller than c for the transformation to make sense.

A Lorentz boost, which is simply a Lorentz transformation that doesn't involve rotation, in direction $\boldsymbol{n}$ with magnitude v

$$t' = \gamma(t - v\boldsymbol{n} \cdot x)$$
$$x' = x + (\gamma - 1)(x \cdot n)n - \gamma tv\boldsymbol{n}$$

Assume two inertial reference frames (t, x, y, z) and (t', x', y', z') and two points P_1 and P_2, the Lorentz group is the set of all transformations between the two reference frames that preserve the speed of light propagating between the two points.

$$c^2(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

In matrix form these are all the linear transformations $\Lambda^T\eta\Lambda = \eta$

$$X' = \begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix}, \qquad \eta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad X = \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}$$

The spacetime interval takes the form

$$X \cdot X = X^T \eta X = X'^T \eta X'$$

And is invariant under the Lorentz transformation

$$X' = \Lambda X$$

Where Λ is a square matrix which can depend on parameters.

i. Light

I'm not sure what light is. But I know I can see something over there. I know what a nice sunset looks like. I know what it feels like to stumble in darkness. Something is over there, and I am here. The relationship between those two things, and really that statement itself, is a definition of light, the photon, the emission and absorption of a quanta of energy between two particles. But how does the universe choose which particles to interact with which where and when. And what too.

i. Gauss's Law (for electricity)

In its integral form, it states that the flux of the electric field out of an arbitrary closed surface is proportional to the electric charge enclosed by the surface, irrespective of how that charge is distributed. Even though the law alone is insufficient to determine the electric field across a surface enclosing any charge distribution, this may be possible in cases where symmetry mandates uniformity of the field. Where no such symmetry exists, Gauss's law can be used in its differential form, which states that the divergence of the electric field is proportional to the local density of charge.

Gauss's law may be expressed

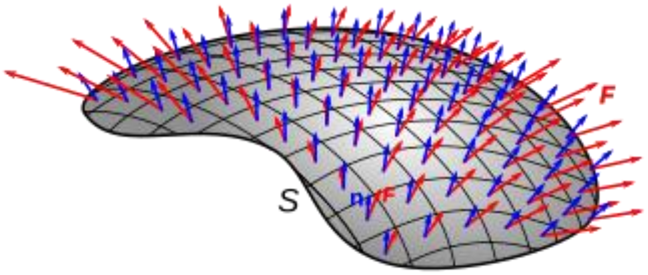
$$\Phi_E = \frac{Q}{\epsilon_0}$$

Where Φ_E is the electric flux through a closed surface S, enclosing any volume V, Q is the total charge enclosed within V, and ϵ_0 is the permittivity of free space.

$$\Phi_E = \oiint_S E \cdot dA$$

ii.i) Flux

Flux describes any effect that appears to pass or travel through a surface.



First the flux at a single point is given

$$j = \frac{I}{A} = \frac{\frac{dq}{dt}}{A}$$

Where j is the given flux, q is the physical quantity that flows, t is time and A is area, that is how much does the quantity flow over time over across an area.

Flux as a scalar field defined a function of a set of point on a surface

$$j(p) = \frac{I(p)}{A(p)}$$

$$I(A,p) = \frac{dq}{dt}(A,p)$$

As before, the surface is assumed to be flat, and the flow is assumed to be everywhere perpendicular to it. However, the flow need not be constant. q is now a function of p, a point on the surface, and A, an area. Rather than measure the total flow through the surface, q measures the flow through the disk with area A centered at p along the surface.

Finally, flux as a vector field:

$$j(p) = \frac{I(p)}{A(p)}$$

$$I(A,p) = \text{BiggestVector}\left(\hat{n}_p \frac{dq}{dt}(A,p,\hat{n})\right)$$

In this case, there is no fixed surface we are measuring over. q is a function of a point, an area, and a direction (given by a unit vector $\hat{n}$), and measures the flow through the disk of area A perpendicular to that unit vector. I is defined picking the unit vector that maximizes the flow around the point, because the true flow is maximized across the disk that is perpendicular to it. The unit vector thus uniquely maximizes the function when it points in the "true direction" of the flow.

If the flux j passes through the area at an angle θ to the area normal $\hat{n}$, the the dot product

$$j \cdot \hat{n} = j \cos\theta$$

For vector flux the surface integral of j over a surface S, gives the proper flowing per unit through the surface

$$\frac{dq}{dt} = \iint_S j \cdot \hat{n} dA$$

ii.ii) Inductance

Inductance is the tendency of an electrical conductor to oppose a change in the electric current flowing through it. The electric current produces a magnetic field around the conductor. The magnetic field strength depends on the magnitude of the electric current and follows any changes in the magnitude of the current. From Faraday's law of induction, any change in magnetic field through a circuit induces an electromotive force (EMF) (voltage) in the conductors, a process known as electromagnetic induction.

Inductance is measured in Henry's that have the SI base units

$$H = \frac{kg \cdot m^2}{s^2 \cdot A^2}$$

Where **A** is ampere, One ampere is equal to 1 coulomb (C) moving past a point per second and is defined in terms of the elementary charge e, at about $6.241509 \times 10^{18} e$ The elementary charge, usually denoted by e, is a fundamental physical constant, defined as the electric charge carried by a single proton or, equivalently, the magnitude of the negative electric charge carried by a single electron, which has charge $-1\ e$

ii.iii) Ampere’s Law and the permittivity of free space.

Two thin, straight, stationary, parallel wires, a distance r apart in free space, each carrying a current I, will exert a force on each other. Ampère's force law states that the magnetic force Fm per length L is given by

$$2 \times 10^{-7} \frac{\left(\frac{kg \cdot m}{s^2}\right)}{m} = \frac{F_m}{L} = \frac{\mu_0}{2\pi}\frac{A^2}{m}$$

The vacuum magnetic permeability is given by,

$$\mu_0 = 4\pi \times 10^{-7} \frac{H}{m} = 1.256 \ldots \times 10^{-6}\ kg \cdot\ m \cdot s^{-2} \cdot A^{-2}$$

$$\epsilon_0 = \frac{1}{\mu_0 c^2}$$

By the divergence theorem, Gauss’s law can also be written in the differential form

$$\nabla \cdot E = \frac{\rho}{\epsilon_0}$$

Where $\nabla \cdot E$ is the divergence of the electric field, and ρ is the charge density, or charge per unit volume.

ii.iv) Divergence

In cartesian coordinates the divergence of a continuously differentiable vector field $F = F_x i + F_y j + F_z k$ is

$$div\ F = \nabla \cdot F = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot \left(F_x, F_y, F_z\right) = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

Vector fields are often illustrated using the example of the velocity field of a fluid, such as a gas or liquid. A moving liquid has a velocity—a speed and a direction—at each point, which can be represented by a vector, so that the velocity of the liquid at any moment forms a vector field. Consider an imaginary closed surface S inside a body of liquid, enclosing a volume of liquid. The flux of liquid out of the volume at any time is equal to the volume rate of fluid crossing this surface, i.e., the surface integral of the velocity over the surface.

Since liquids are incompressible, the amount of liquid inside a closed volume is constant; if there are no sources or sinks inside the volume then the flux of liquid out of S is zero. If the liquid is moving, it may flow into the volume at some points on the surface S and out of the volume at other points, but the amounts flowing in and out at any moment are equal, so the net flux of liquid out of the volume is zero.

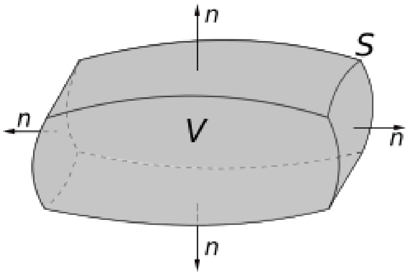
However if a source of liquid is inside the closed surface, such as a pipe through which liquid is introduced, the additional liquid will exert pressure on the surrounding liquid, causing an outward flow in all directions. This will cause a net outward flow through the surface S. The flux outward through S equals the volume rate of flow of fluid into S from the pipe. Similarly, if there is a sink or drain inside S, such as a pipe which drains the liquid off, the external pressure of the liquid will cause a velocity throughout the liquid directed inward toward the location of the drain. The volume rate of flow of liquid inward through the surface S equals the rate of liquid removed by the sink.

If there are multiple sources and sinks of liquid inside S, the flux through the surface can be calculated by adding up the volume rate of liquid added by the sources and subtracting the rate of liquid drained off by the sinks. The volume rate of flow of liquid through a source or sink (with the flow through a sink given a negative sign) is equal to the divergence of the velocity field at the pipe mouth, so adding up (integrating) the divergence of the liquid throughout the volume enclosed by S equals the volume rate of flux through S. This is the divergence theorem.

The divergence theorem is employed in any conservation law which states that the total volume of all sinks and sources, that is the volume integral of the divergence, is equal to the net flow across the volume's boundary.

Suppose V is a subset of $\mathbb{R}^n$ (in the case n = 3, V represents a volume in three-dimensional space) which is **compact** and has a smooth boundary S. If F is a continuously differentiable vector field defined on a neighbourhood of V then,

$$\iiint\limits_V (\nabla \cdot F) dV = \oiint_S (F \cdot \hat{n}) dS$$



The integral form of Gauss's law is:

$$\oiint_S E \cdot dA = \frac{Q}{\epsilon_0}$$

for any closed surface S containing charge Q. By the divergence theorem, this equation is equivalent to:

$$\iiint\limits_V \nabla \cdot E dV = \frac{Q}{\epsilon_0}$$

for any volume V containing charge Q. By the relation between charge and charge density, this equation is equivalent to:

$$\iiint\limits_V \nabla \cdot E dV = \iiint\limits_V \frac{\rho}{\epsilon_0} V$$

for any volume V. In order for this equation to be simultaneously true for every possible volume V, it is necessary (and sufficient) for the integrands to be equal everywhere. Therefore, this equation is equivalent to:

$$\nabla \cdot E = \frac{\rho}{\epsilon_0}$$

The is the first of maxwells equations.

ii. Gauss’s Law for Magnetism or the absence of free magnetic poles.

The differential form for Gauss's law for magnetism is:

$$\nabla \cdot B = 0$$

where B denotes the magnetic field. The integral form of Gauss's law for magnetism states

$$\Phi_B = \oiint_S B \cdot dS = 0$$

The law in this form states that for each volume element in space, there are exactly the same number of "magnetic field lines" entering and exiting the volume. No total "magnetic charge" can build up in any point in space. For example, the south pole of the magnet is exactly as strong as the north pole, and free-floating south poles without accompanying north poles (magnetic monopoles) are not allowed. In contrast, this is not true for other fields such as electric fields or gravitational fields, where total electric charge or mass can build up in a volume of space. It is equivalent to the statement that magnetic monopoles do not exist.

The magnetic field B can be depicted via field lines (also called flux lines) – that is, a set of curves whose direction corresponds to the direction of B, and whose areal density is proportional to the magnitude of B. Gauss's law for magnetism is equivalent to the statement that the field lines have neither a beginning nor an end: Each one either forms a closed loop, winds around forever without ever quite joining back up to itself exactly, or extends to infinity.

iii. Faraday's Law (of Inductance)

The most widespread version of Faraday's law states:

The electromotive force around a closed path is equal to the negative of the time rate of change of the magnetic flux enclosed by the path.

For a loop of wire in a magnetic field, the magnetic flux Φ_B is defined for any surface Σ whose boundary is the given loop. Since the wire loop may be moving, we write $\Sigma(t)$ for the surface, The magnetic flux is the surface integral,

$$\Phi_B = \oiint_{\Sigma(t)} B(t) \cdot dA$$

iii.i) The Maxwell-Faraday Equation

The Maxwell–Faraday equation states that a time-varying magnetic field always accompanies a spatially varying (also possibly time-varying), non-conservative electric field, and vice versa. The Maxwell–Faraday equation is

$$\nabla \times E = -\frac{\partial B}{\partial t}$$

iii.ii) Curl

A vector field whose curl is zero is called irrotational.

In three dimensions the curl is given by

$$\nabla \times F = \begin{bmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$

$$\nabla \times F = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\hat{\imath} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\hat{\jmath} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\hat{k}$$

The curl of a vector field denoted $\nabla \times F$, is an operator that maps k finitely differentiable smooth functions to $k - 1$ finitely differentiable smooth functions. Unlike the gradient and divergence, curl as formulated in vector calculus does not generalize simply to other dimensions; some generalizations are possible, but only in three dimensions is the geometrically defined curl of a vector field again a vector field. This deficiency is a direct consequence of the limitations of vector calculus; on the other hand, when expressed as an antisymmetric tensor field via the wedge operator of geometric calculus, the curl generalizes to all dimensions.

iv. Ampere’s circuital law law with Maxwell’s addition.

In 1820 Danish physicist Hans Christian Ørsted discovered that an electric current creates a magnetic field around it, when he noticed that the needle of a compass next to a wire carrying current turned so that the needle was perpendicular to the wire. He investigated and discovered the rules which govern the field around a straight current-carrying wire:

The magnetic field lines encircle the current-carrying wire.
The magnetic field lines lie in a plane perpendicular to the wire.
If the direction of the current is reversed, the direction of the magnetic field reverses.
The strength of the field is directly proportional to the magnitude of the current.
The strength of the field at any point is inversely proportional to the distance of the point from the wire.

This sparked a great deal of research into the relation between electricity and magnetism. André-Marie Ampère investigated the magnetic force between two current-carrying wires, discovering Ampère's force law. In the 1850s Scottish mathematical physicist James Clerk Maxwell generalized these results and others into a single mathematical law. The original form of Maxwell's circuital law, which he derived as early as 1855 in his paper "On Faraday's Lines of Force"[9] based on an analogy to hydrodynamics, relates magnetic fields to electric currents that produce them. It determines the magnetic field associated with a given current, or the current associated with a given magnetic field.

The original circuital law only applies to a magnetostatic situation, to continuous steady currents flowing in a closed circuit. For systems with electric fields that change over time, the original law (as given in this section) must be modified to include a term known as Maxwell's correction.

The integral form of the original circuital law is a line integral of the magnetic field around some closed curve C (arbitrary but must be closed). The curve C in turn bounds both a surface S which the electric current passes through (again arbitrary but not closed—since no three-dimensional volume is enclosed by S) and encloses the current. The mathematical statement of the law is a relation between the circulation of the magnetic field around some path (line integral) due to the current which passes through that enclosed path (surface integral).

In terms of total current, (which is the sum of both free current and bound current) the line integral of the magnetic B-field (in teslas, T) around closed curve C is proportional to the total current I_{enc} passing through a surface S (enclosed by C). In terms of free current, the line integral of the magnetic H-field (in amperes per metre, $A \cdot m^{-1}$) around closed curve C equals the free current $I_{f,enc}$ through a surface S.

Using B-field and total current

The Integral form is given,

$$\oint_C B \cdot dl = \mu_0 \oiint_S J \cdot dS = \mu_0 I_{enc}$$

The differential form is given

$$\nabla \times B = \mu_0 J$$

J is the total current density (in amperes per square meter. $A \cdot m^{-2}$)
$\oint_C$ is the closed line integral around closed curve C
$\oiint_S$ denotes a surface integral over the surface S bounded by the curve C
dl is an infinitesimal element (a differential) of the curve C (i.e. a vector with magnitude equal to the length of the infinitesimal line element, and direction given by the tangent to the curve C)
dS is the vector area of an infinitesimal element of surface S (that is, a vector with magnitude equal to the area of the infinitesimal surface element, and direction normal to surface S. The direction of the normal must correspond with the orientation of C by the right-hand rule), see below for further explanation of the curve C and surface S.

There are two important issues regarding the circuital law that require closer scrutiny. First, there is an issue regarding the continuity equation for electrical charge. In vector calculus, the identity for the divergence of a curl states that the divergence of the curl of a vector field must always be zero. Hence,

$$\nabla \cdot (\nabla \times B) = 0$$

And so Ampere's circuital law implies,

$$\nabla \cdot J = 0$$

i.e. that current density is solenoidal (solenoidal by definition means $\nabla \cdot v = 0$.)
But in general, to whatever it means that reality follows the continuity equation for electric charge is given,

$$\nabla \cdot J = -\frac{\partial \rho}{\partial t}$$

Which is nonzero for a time varying charge density.

iv.i) Continuity equations

A continuity equation or transport equation is an equation that describes the transport of some quantity. It is particularly simple and powerful when applied to a conserved quantity, but it can be generalized to apply to any extensive quantity. Since mass, energy, momentum, electric charge and other natural quantities are conserved under their respective appropriate conditions, a variety of physical phenomena may be described using continuity equations.

iv.i.i) With regards to flux

A continuity equation is useful when a flux can be defined.

e.g. electric flux

$$\Phi_E = \oiint_S E \cdot dA$$

And more generally,

$$Flux\ (the\ rate\ q\ flowing\ through\ the\ imaginary\ surface\ S) = \iint_S j \cdot dS$$

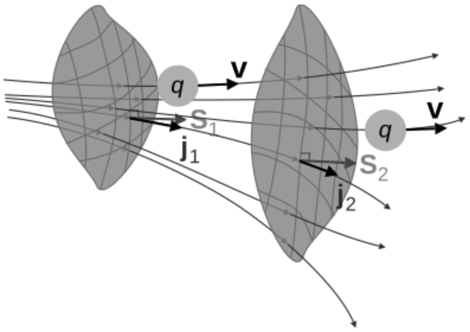


Figure 1 Illustration of how the fluxes, or flux densities, j_1 and j_2 of a quantity q pass through open surfaces S_1 and S_2. (vectors $S_1$and S_2represent vector areas that can be differentiated into infinitesimal area elements).

The integral form of the continuity equation states that:

> The amount of q in a region increases when additional q flows inward through the surface of the region, and decreases when it flows outward; The amount of q in a region increases when new q is created inside the region, and decreases when q is destroyed;

That is,

$$\frac{dq}{dt} + \oint_S j \cdot dS = \Sigma$$

Σ is the net rate q is being generated inside a volume V per unit time when q is being it is called a source of q, and it make Σ positive, when q is being destroyed it is called a sink of q, and it makes Σ more negative. This term is sometimes written $dq/dt|_{gen}$ or the total charge of q from it generation or destruction inside the control volume.
By the divergence theorem, a general continuity equation can also be written in differential form,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j = \sigma$$

Regarding the issue the propagation of electromagnetic waves, in free space where,

$$J = 0$$

The circuital law implies

$$\nabla \times B = 0$$

i.e. that the magnetic field is irrotational but to maintain consistency with the continuity equation for electric charge we must have,

$$\nabla \times B = \frac{1}{c^2}\frac{\partial E}{\partial t}$$

To treat these situations, the contribution of displacement current must be added to the current term in the circuital law.

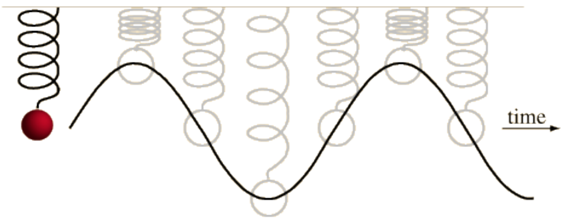
James Clerk Maxwell conceived of displacement current as a polarization current in the dielectric vortex sea, which he used to model the magnetic field hydrodynamically and mechanically.[17] He added this displacement current to Ampère's circuital law at equation 112 in his 1861 paper "On Physical Lines of Force".

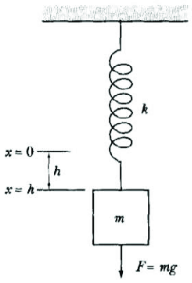
Displacement Current

In free space, the displacement current is related to the time rate of change of electric field.

The Model

Given a simple harmonic oscillator such as a spring:





The work done by a conservative force, such in the case of the spring, the gravitational force, is

$$\mathrm{W} = -\Delta U$$

where ΔU is the change in the potential energy associated with the force.

Kinetic Energy + Potential Energy = Total Energy

$$K.E + P.E = E$$

The change in kinetic energy is equal to the work done associated with the force and displacement by newtons second law of motion.

$$\Delta K = W$$

$$\Delta K = \int F\, dx = \int ma\, dx = m \int \frac{dv}{dt}\, dx = m \int \frac{dx}{dt}\, dv = m \int v\, dv$$

So we consider the change in kinetic energy, to be the force exerted at each point along some displacement, that is proportional to its mass and acceleration, proportional to its momentum and its velocity.

$$\Delta K = m \int_{v_0}^{v} v dv = \frac{1}{2} m v^2 - \frac{1}{2} m v_0^2$$

Naturally, the kinetic energy of an object at rest should be zero. Thus, an object's kinetic energy is defined mathematically by the following equation…

$$\mathrm{K} = \frac{1}{2} m\, v^2$$

Classically Kinetic Energy

$$\mathrm{K.E} = \frac{1}{2} m\, v^2 = \frac{p^2}{2m}$$

Classically Potential Energy follows in the same fashion

$$\mathrm{P.E} = \int F \cdot dx$$

$$W = \int_0^t F \cdot vdt = -\int_0^t kxv_x dt = -\int_0^t kx\frac{dx}{dt}dt = \int_{x(t_0)}^{x(t)} kxdx = \frac{1}{2}kx^2$$

$$P.E = \frac{1}{2}kx^2$$

$$\frac{p^2}{2m} + \frac{1}{2}kx^2 = E$$

Momentum in quantum mechanics is given by

$$p = \frac{\hbar}{i}\frac{\partial}{\partial x}$$

Swapping this into our equation gives us

$$\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}kx^2$$

We will have any arbitrary potential U(x)

$$H = \left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + U(x)\right)$$

We introduce the wave function $\psi(x,t)$

The Hamiltonian form of the Schrodinger equation is given:

$$H\psi(x,t) = E\psi(x,t)$$

That is,

$$\left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + P.E(x,t)\right)\psi(x,t) = E\psi(x,t)$$

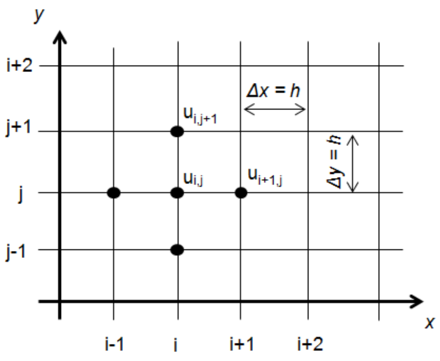
$$(Kinetic\ Energy + Potential\ Energy)ParticleState = Energy \times ParticleState$$

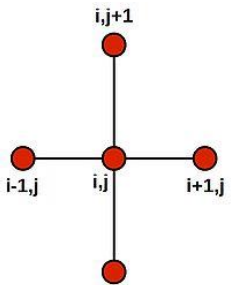
$$\frac{\partial^2}{\partial x^2} = \nabla^2 f(x)$$

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

$$\frac{\partial f}{\partial x}(i\ ,j) \approx f(i+1,j) - f(i,j)$$

$$\frac{\partial f}{\partial y}(i,j) \approx f(i,j+1) - f(i,j)$$





$$\nabla f(x_i, y_i) = \left(\frac{\partial f}{\partial x}(x_i\,, y_j), \frac{\partial y}{\partial x}(x_i\,, y_j)\right)$$

Which becomes.

$$\nabla f(x_i, y_i) = \big(f(i+1,j) - f(i,j), f(i,j+1) - f(i,j)\big)$$
$$\frac{\partial^2 f}{\partial x^2} = \big(f(i+1,j) - f(i,j)\big) - \big(f(i,j) - f(i-1,j)\big)$$

Which is,

$$\frac{\partial^2 f}{\partial x^2} = f(i+1,j) + f(i-1,j) - 2f(i,j)$$
$$\frac{\partial^2 f}{\partial y} = f(i,j+1) + f(i,j-1) - 2f(i,j)$$

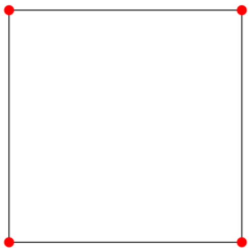
Thus,

$$\nabla^2 f = f(i+1,j) + f(i-1,j) + f(i,j+1) + f(i,j-1) - 4f(i,j)$$

Say we have a 2x2 space. At each point in space there is a complex number. There are 2^2 *states.*

$$\psi_{xy} = \begin{pmatrix} \psi_{00} & \psi_{01} \\ \psi_{10} & \psi_{11} \end{pmatrix}$$

$$\psi_{xy} = \begin{pmatrix} \psi_{00} \\ \psi_{01} \\ \psi_{10} \\ \psi_{11} \end{pmatrix}$$



$$\nabla^2 \psi_{xy} = \begin{pmatrix} \nabla^2\psi_{00} \\ \nabla^2\psi_{01} \\ \nabla^2\psi_{10} \\ \nabla^2\psi_{11} \end{pmatrix} = \begin{pmatrix} \psi_{01} + \psi_{10} - 4\psi_{00} \\ \psi_{11} - 4\psi_{01} \\ \psi_{11} - 4\psi_{10} \\ -4\psi_{11} \end{pmatrix} = \begin{pmatrix} -4 & 1 & 1 & 0 \\ 1 & -4 & 0 & 1 \\ 1 & 0 & -4 & 1 \\ 0 & 1 & 1 & -4 \end{pmatrix} \begin{pmatrix} \psi_{00} \\ \psi_{01} \\ \psi_{10} \\ \psi_{11} \end{pmatrix}$$

$$\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -4 & 1 & 1 & 0 \\ 1 & -4 & 0 & 1 \\ 1 & 0 & -4 & 1 \\ 0 & 1 & 1 & -4 \end{pmatrix}$$

If A is $n \times\ n$, B is $m \times m$ and I_k denotes the $k \times k$ identity matrix, then we can define a “Kronecker sum” $\oplus$

$$A \oplus B = A \otimes I_m + I_n \otimes B$$

The Kronecker product of matrices corresponds to the abstract tensor product of linear maps. Specifically, if the vector spaces V, W, X and Y have bases $\{v_1, \dots, v_m\}, \{w_1, \dots, 2_n\}, \{x_1, \dots x_d\}\ and\ \{y_1, \dots, y_e\}$, respectively, and if the matrices A and B present linear transformations $S : V \longrightarrow X$ and $T : W \ \longrightarrow Y$, respectively in the appropriate bases, then the matrix $A \otimes B$ represents the tensor product of two maps, $S \otimes T :\ V \otimes W \longrightarrow X \otimes Y$ with respect to the basis $\{v_1 \otimes w_1, v_1 \otimes w_2, \dots v_2 \otimes w_1, \dots, v_m \otimes w_n\}\ of\ V \otimes W$ and the similarly defined basis of $X \otimes Y$ with the property that $A \otimes B\big(v_i \otimes w_j\big) = (Av_i) \otimes \big(Bw_j\big)$, where i and j are integers in the proper range.

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

$$e.g. \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} 1\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 2\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \\ 3\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 4\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 5 & 0 & 10 \\ 6 & 7 & 12 & 14 \\ 0 & 15 & 0 & 20 \\ 18 & 21 & 24 & 28 \end{bmatrix}$$

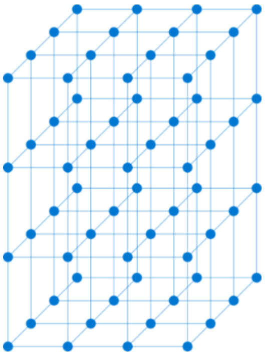
It is going to be our attempt to solve the linear equation:

$$A\psi(x, t = i + 1) = B\psi(x, t = i)$$

First, we will introduce a complex valued wave function. That is

$$s \in \psi(\mathrm{x}, \mathrm{t}), \qquad s \in \mathbb{C}$$

Our space is some finite subset of $\mathbb{Z}^3$ $\boldsymbol{x} = (x, y, z) \mid x, y, z \in \mathbb{Z}^3$. Our space has a length L in each direction, for simplicity we consider a square space where $-\frac{L}{2} < x < \frac{L}{2}$. There are $N \times N \times N$ points.



Where Every point is a complex number. We recall the isomorphism of the structure of $\mathbb{C}$ to $\mathbb{R}^2$

We have two real spaces that are independent.

We will start with a simple gaussian distribution for our particle state $\psi(x, t = 0)$

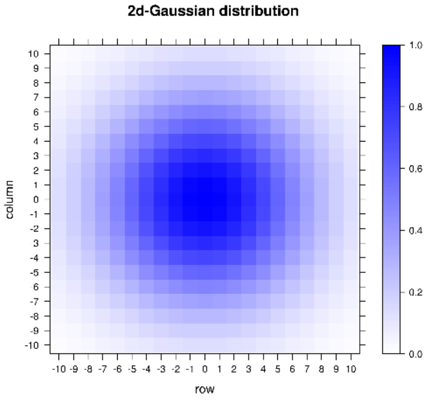
$$\psi(x, 0) = e^{-\frac{1}{2}x^2}$$

We can modify the standard deviation σ and centre μ of the gaussian distribution

$$\psi(x, 0) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Position Space

Momentum Space



We add a complex part, we introduce phase velocity into the gaussian, this part exists in the momentum space.

$$\psi(x,0) = e^{-\frac{1}{2}x^2} \cdot e^{-ip_0 x}$$

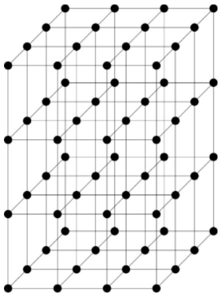
For $\psi(x, t = i + 1)$, from an initial wave function $\psi(x, t = 0)$,

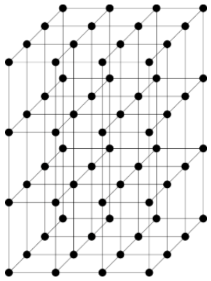
Given,

$$H = K.E + P.E$$

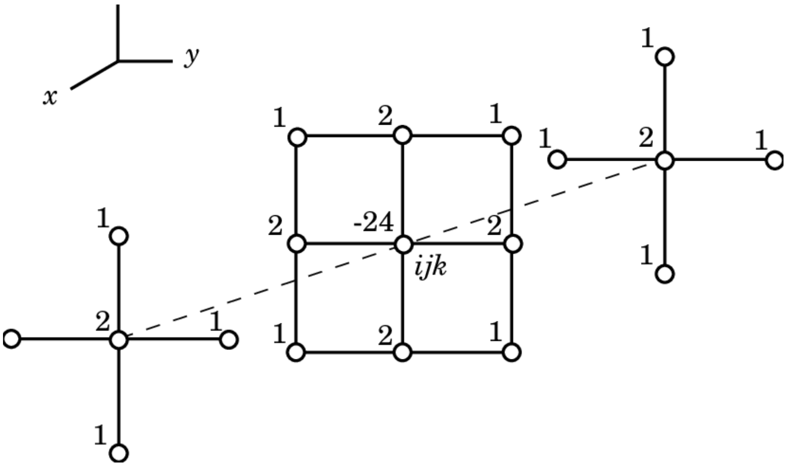
$$H\,\psi(x, t = i) = E\,\psi(x, t = i)$$

$$A\,\psi(x, t = i + 1) = B\,\psi(x, t = i)$$





There is an $N \times N$ tridiagonal matrix T, The matrix T represents the discrete Laplacian operator, a central difference approximation to the second derivative.



$$T = -\frac{1}{mdx^2}\frac{\hbar^2}{2m}\begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & 1 & -2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & -2 \end{bmatrix} = \begin{bmatrix} \frac{2}{mdx^2}\frac{\hbar^2}{2m} & -\frac{1}{mdx^2}\frac{\hbar^2}{2m} & 0 & \cdots & 0 \\ -\frac{1}{mdx^2}\frac{\hbar^2}{2m} & \frac{2}{mdx^2}\frac{\hbar^2}{2m} & -\frac{1}{mdx^2}\frac{\hbar^2}{2m} & \ddots & \vdots \\ 0 & -\frac{1}{mdx^2}\frac{\hbar^2}{2m} & \frac{2}{mdx^2}\frac{\hbar^2}{2m} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\frac{1}{mdx^2}\frac{\hbar^2}{2m} \\ 0 & \cdots & 0 & -\frac{1}{mdx^2}\frac{\hbar^2}{2m} & \frac{2}{mdx^2}\frac{\hbar^2}{2m} \end{bmatrix}$$

This matrix is then scaled by -— $\frac{1}{mdx^2}\frac{\hbar^2}{2m}$, which represents the kinetic energy term in the finite difference approximation.

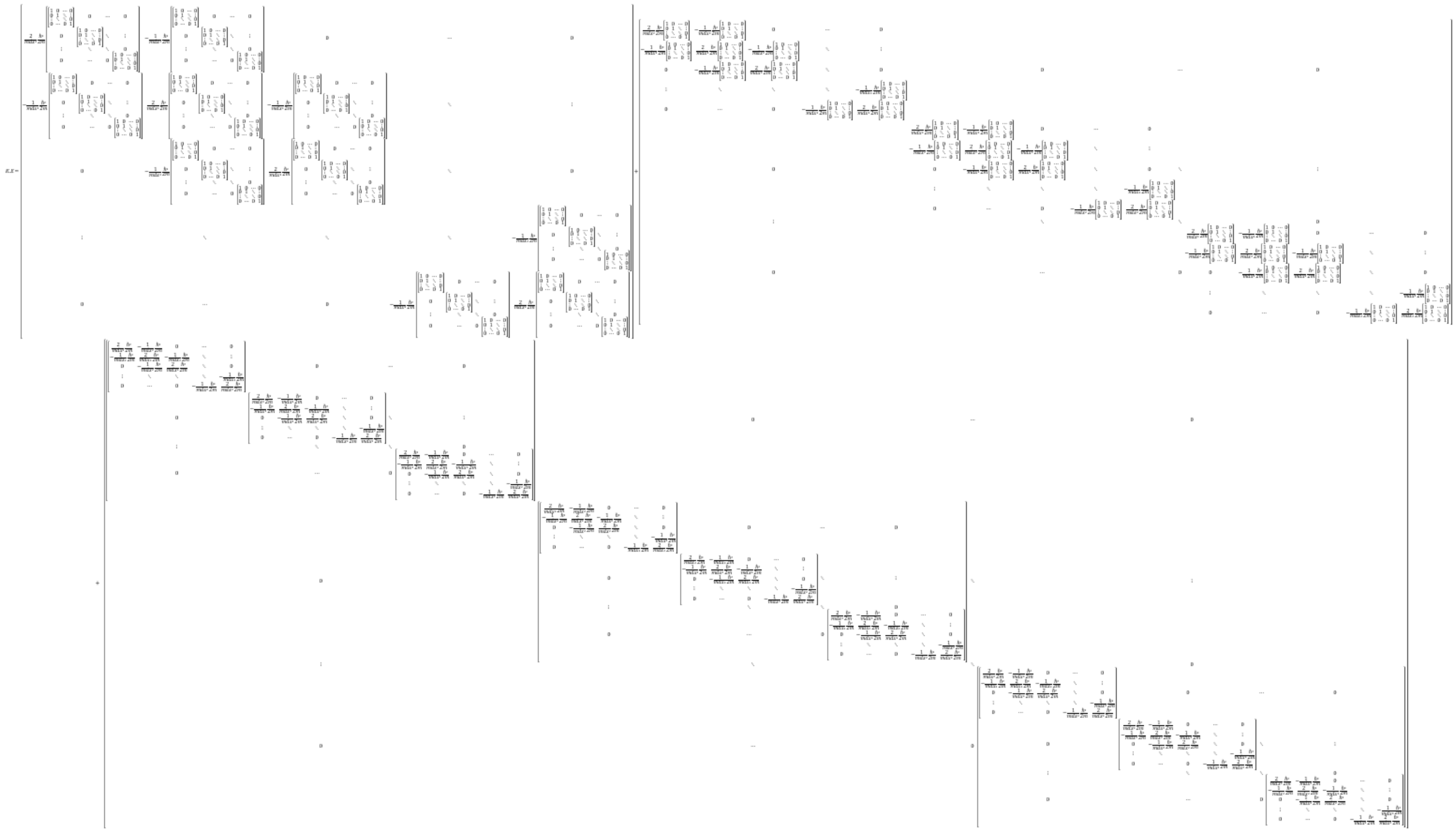
We use Kronecker products to extend the 1D kinetic operator T to 3D space.

$$K.E = T \otimes (I_N \otimes I_N) + I_N \otimes (T \otimes I_N) + I_N \otimes (I_N \otimes T)$$

$$(I_N \otimes I_N) = \begin{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} & 0 & \cdots & 0 \\ 0 & \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \end{bmatrix}$$

$$T \otimes (I_N \otimes I_N) = \begin{bmatrix} \frac{2}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & -\frac{1}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & 0 & \cdots & 0 \\ -\frac{1}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & \frac{2}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & -\frac{1}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & \ddots & \vdots \\ 0 & -\frac{1}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & \frac{2}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\frac{1}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) \\ 0 & \cdots & 0 & -\frac{1}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) & \frac{2}{mdx^2}\frac{\hbar^2}{2m}(I_N \otimes I_N) \end{bmatrix}$$

The Almost Full Expansion of the Kinetic Matrix K.E = $T \otimes (I_N \otimes I_N) + I_N \otimes (T \otimes I_N) + I_N \otimes (I_N \otimes T)$



For every i, j and k position in our space we have some potential,

$$U(i,j,k) = \left[\begin{bmatrix} a_{00} & \cdots & a_{i0} \\ \vdots & \ddots & \vdots \\ a_{0j} & \cdots & a_{ij} \end{bmatrix}_0 \begin{bmatrix} a_{00} & \cdots & a_{i0} \\ \vdots & \ddots & \vdots \\ a_{0j} & \cdots & a_{ij} \end{bmatrix}_1 \cdots \begin{bmatrix} a_{00} & \cdots & a_{i0} \\ \vdots & \ddots & \vdots \\ a_{0j} & \cdots & a_{ij} \end{bmatrix}_k\right]$$

For every i, j and k in N there is an element of the potential U(i,j,k) and w *such that* $w = i \cdot N \cdot N + j \cdot N + k$

We have the diagonal matrix

$$P.E = \begin{bmatrix} U_0 & 0 & \cdots & 0 \\ 0 & U_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_w \end{bmatrix}$$

Where each diagonal entry is an element of our 3D potential.

$$H = \left(T \otimes (I_N \otimes I_N) + I_N \otimes (T \otimes I_N) + I_N \otimes (I_N \otimes T)\right) + \begin{bmatrix} U_0 & 0 & \cdots & 0 \\ 0 & U_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_w \end{bmatrix}$$

The full size of H is $N^3 \times N^3$

$$B = I_H - \frac{1}{2}\frac{i\,dt}{\hbar}H$$

$$A = I_H + \frac{1}{2}\frac{i\,dt}{\hbar}H$$

We are solving for x

$$\boldsymbol{Ax} = \boldsymbol{b}$$

Where,

$$\boldsymbol{b} = B \cdot \psi(t = 0)$$

$$\boldsymbol{b} = \left(I_H - \frac{1}{2}\frac{i\,dt}{\hbar}H\right)\psi(t = 0)$$

$$\boldsymbol{b} = \left(I_H - \frac{1}{2}\frac{i\,dt}{\hbar}\left(\left(T \otimes (I_N \otimes I_N) + I_N \otimes (T \otimes I_N) + I_N \otimes (I_N \otimes T)\right) + \begin{bmatrix} U_0 & 0 & \cdots & 0 \\ 0 & U_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_w \end{bmatrix}\right)\right)\psi(t = 0)$$

$$\boldsymbol{b} = \left(I_H - \frac{1}{2}\frac{i\,dt}{\hbar}\left(\left(T \otimes (I_N \otimes I_N) + I_N \otimes (T \otimes I_N) + I_N \otimes (I_N \otimes T)\right) + \begin{bmatrix} U_0 & 0 & \cdots & 0 \\ 0 & U_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_w \end{bmatrix}\right)\right)\left(e^{-\frac{1}{2}\hat{x}^2} \cdot e^{-ip_0\hat{x}}\right)$$

And

$$\boldsymbol{A} = I_H + \frac{1}{2}\frac{i\,dt}{\hbar}\left(\left(T \otimes (I_N \otimes I_N) + I_N \otimes (T \otimes I_N) + I_N \otimes (I_N \otimes T)\right) + \begin{bmatrix} U_0 & 0 & \cdots & 0 \\ 0 & U_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_w \end{bmatrix}\right)$$

We use an iterative method to solve such as SciPy’s GCROT

$$\psi(t + 1) = x = GCROT(A, b)$$

We will first explore an algebraically similar method. The conjugate residual method.

The conjugate gradient method is an algorithm for the numerical solution of systems of linear equations.

Suppose we want to solve the systems of linear equations

$$Ax = b$$

For the vector x. where the known $n \times n$ matrix A is symmetric $i.e\ A^T = A$, positive-definite ($i.e\ x^T Ax > 0$ for all nonzero vector $x \in \mathbb{R}^n$) and real, and b is known as well. We denote the unique solution of this system x_*.

We say that two non-zero vectors u and v are conjugate (with respect to A) if

$$u^T Av = 0$$

Since A is symmetric and positive definite, the left-hand side defines an inner product,

$$u^T Av = \langle u, v\rangle_A := \langle Au, v\rangle = \langle u, A^T v\rangle = \langle u, Av\rangle$$

An inner product is a generalization of the dot product. In a vector space, it is a way to multiply vectors together, with the result of this multiplication being a scalar.

A function $\langle .,. \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is an inner product on $\mathbb{R}^n$ if an only if there exists a symmetric positive-definite matrix M such that $\langle x, y\rangle = x^T My$ for all $x, y \in \mathbb{R}^n$. If M is the identity matrix, then $\langle x, y\rangle = x^T My$ is the dot product.

For $\mathbb{R}^2$, where M is positive-definite (which happens if an only if $\det M = ad - b^2 > 0$ and one/both diagonal elements are positive) then for any $x := [x_1, x_2]^T, y := [y_1, y_2]^T \in \mathbb{R}^2$,

$$\langle x, y\rangle := x^T My = [x_1 x_2]\begin{bmatrix} a & b \\ b & d \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = ax_1y_1 + bx_1y_2 + bx_2y_1 + dx_2y_2$$

Every inner product space on $\mathbb{R}^2$ is of this form (where $b \in \mathbb{R}, a > 0\ and\ d > 0\ satisfy\ ad > b^2$)

The general form of an inner product on $\mathbb{C}^n$ is known as the Hermitian form and is given by:

$$\langle x, y\rangle = y^\dagger Mx = \overline{x^\dagger My}$$

The inner product of a pair of complex numbers z and w is the product of z with the complex conjugate of w:

$$\langle z, w\rangle = z\overline{w}$$

A second example is the space $\mathbb{C}^2$ whose elements are pairs of complex numbers $z = (z_1, z_2)$. Then the inner product of z with another such vector $w = (w_1, w_2)$ is given by,

$$\langle z, w\rangle = z_1\overline{w_1} + z_2\overline{w_2}$$

A Hilbert space is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product. To say that a complex vector space H is a complex inner product space means that there is an inner product $\langle x, y\rangle$ associating a complex number to each pair of elements x, y that satisfies the following properties,

1. The inner product is conjugate symmetric; that is, the inner product of a pair of elements is equal to the complex conjugate of the inner product of the swapped elements.
$$\langle y, x\rangle = \overline{\langle x, y\rangle},$$
Importantly this implies that $\langle x, x\rangle$ is a real number.

2. The inner product is linear in its first argument. For all complex numbers a and b,
$$\langle ax_1 + bx_2, y\rangle = a\langle x_1, y\rangle + b\langle x_2, y\rangle$$

3. The inner product of an element with itself is positive-definite:
$$\langle x, x\rangle > 0 \quad if\ x \neq 0$$
$$\langle x, x\rangle = 0 \quad if\ x = 0$$

It follows from 1 and 2 that a complex inner product is antilinear (conjugate linear) in its second argument, meaning that,

$$\langle x. ay_1 + by_2\rangle = \bar{a}\langle x, y_1\rangle + \bar{b}\langle x, y_2\rangle$$

Suppose V is a vector-space. A function f that converts ordered pairs of vectors into scalars is a bilinear form if it satisfies the rules:

- $f(u+v,w) = f(u,w) + f(v,w)$
- $f(ku,v) = k \cdot f(u,v)$
- $f(u,v+w) = f(u,v) + f(u,w)$
- $f(u,kv) = k \cdot f(u,v)$

If f is a bilinear form, it is symmetric if it satisfies:

- $f(u,v) = f(v,u)$

If f is a symmetric bilinear form, it is positive definite if it satisfies:

- $f(u,u) > 0, where\ u \neq 0$

An inner product is a positive-definite symmetric bilinear form (While a real inner product $\langle x, y \rangle$ is bilinear (linear in each variable separately), there is a slight difference for complex inner products, A complex inner product $\langle x|y \rangle$ is linear in x and is conjugate linear in y).

An inner-product space is a vector space with an inner product. By positive-definiteness, the norm $|v|$ of a vector v in an inner-product space can be defined to be the non-negative square root of $\langle v, v \rangle$, $\|v\| = \sqrt{\langle v, v \rangle}$

The real numbers $\mathbb{R}$ are a vector space over $\mathbb{R}$ that becomes an inner product space with arithmetic multiplication as its inner product,

$$\langle x. y \rangle \coloneqq xy \quad for\ x, y \in \mathbb{R}$$

The complex numbers $\mathbb{C}$ are a vector space over $\mathbb{C}$ that becomes an inner product space with the inner product,

$$\langle x, y \rangle \coloneqq x\bar{y} \quad for\ x, y \in \mathbb{C}$$

More generally the real n-space $\mathbb{R}^n$ with the dot product is an inner product space, an example of a Euclidean vector space.

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = x^T y = \sum_{i=1}^{n} x_i y_i$$

Two vectors u and v are conjugate if an only if they are orthogonal with respect to the inner product

$$u^T A v = \langle u, v \rangle_A \coloneqq \langle Au, v \rangle = \langle u, A^T v \rangle = \langle u, Av \rangle$$

Being conjugate is a symmetric relation: if u is conjugate to v, then v is conjugate to u.

Suppose that

$$P = \{p_1, \dots, p_n\}$$

Is a set of n mutually conjugate vectors with respect to A, i.e. $p_i^T A p_j = 0$ for all $i \neq j$. Then P forms a basis for $\mathbb{R}^n$, and we may express the solution x_* of $Ax = b$ in this basis:

$$x_* = \sum_{i=1}^{n} a_i p_i \Rightarrow A x_* = \sum_{i=1}^{n} a_i A p_i$$

Left multiplying the problem $Ax = b$ with the vector p_k^T yields,

$$p_k^T b = p_k^T A x_* = \sum_{i=1}^{n} a_i p_k^T A p_i = \sum_{i=1}^{n} a_i \langle p_k, p_i \rangle_A = a_k \langle p_k, p_i \rangle_A$$

And so,

$$a_k = \frac{\langle p_k, b \rangle_A}{\langle p_k, p_k \rangle_A}$$

This gives the following method for solving the equation $Ax = b$: find a sequence of n conjugate directions, and then compute the coefficients a_k.

As an iterative method, if we choose the conjugate vectors p_k carefully, then we may not need all of them to obtain a good approximation to the solution x_*. So, we want to regard the conjugate gradient method as an iterative method. This also allows us to approximately solve systems where n is so large that the direct method would take too much time.

We denote the initial guess for x_* by x_0, we can assume without the loss of generality that $x_0 = 0$. Starting with x_0 we search for the solution and in each iteration, we need a metric to tell us whether we are closer to the solution x_*. This comes from the fact x_* is also the unique minimizer of the following quadratic function

$$f(x) = \frac{1}{2}x^T Ax - x^T b, \qquad x \in \mathbb{R}^n$$

The existence of a unique minimizer is apparent as its Hessian matrix of second derivatives is symmetric positive-definite

$$H\big(f(x)\big) = A$$

And that the minimizer $\nabla f(x)$ solves the initial problem follows from its first derivative,

$$\nabla f(x) = Ax - b$$

This suggests taking the first basis vector p_0 to be the negative of the gradient of f at $x = x_0$. The gradient of f equals $Ax - b$. Starting with an initial guess x_0, this means we take $p_0 = b - Ax_0$. The other vectors in the basis will be conjugate to the gradient, hence the name conjugate gradient method. Note that p_0 is also the residual provided by this initial step of the algorithm.

Let r_k be the residual at the kth step

$$r_k = b - Ax_k$$

As observed above, r_k, is the negative gradient of f at x_k, so the gradient descent method would require to move in the direction of r_k. Hence, however, we insist that the directions p_k must conjugate to each other. A practical way to enforce this is requiring that the next search direction be built out of the current residual and all previous search directions. This conjugate constraint is an orthonormal-type constraint and hence the algorithm can be viewed as an example of a Gram-Schmidt orthonormalization. This gives the following expression:

$$p_k = r_k - \sum_{i<k} \frac{p_i^T A r_k}{p_i^T A p_i} p_i$$

Following this direction, the next optimal location is given by,

$$x_{k+1} = x_k + a_k p_k$$

With

$$a_k = \frac{p_k^T(b - Ax_k)}{p_k^T A p_k} = \frac{p_k^T A r_k}{p_k^T A p_k}$$

Where the last equality follows from the definition of r_k. The expression for a_k can be derived if one substitutes the expression for x_{k+1} into f and minimizing it with respect to a_k

$$f(x_{k+1}) = f(x_k + a_k p_k) \coloneqq g(a_k)$$

$$g'(a_k) \overset{!}{=} 0 \Rightarrow a_k = \frac{p_k^T(b - Ax_k)}{p_k^T A p_k}$$

The above algorithm gives the most straightforward explanation of the conjugate gradient method. Seemingly, the algorithm as stated requires storage of all previous searching directions and residue vectors, as well as many matrix–vector multiplications, and thus can be computationally expensive. However, a closer analysis of the algorithm shows that r_i is orthogonal to r_j , i.e. $r_i^T r_j = 0$, for $i \neq j$. And p_i is A-orthogonal to p_j , i.e. $p_i^T A p_j = 0$, for $i \neq j$ This can be regarded that as the algorithm progresses, p_i and r_i span the same Krylov subspace. Where r_i form the orthogonal basis with respect to the standard inner product, and p_i form the orthogonal basis with respect to the inner product induced by A. Therefore x_k can be regarded as the projection of x on the Krylov subspace.

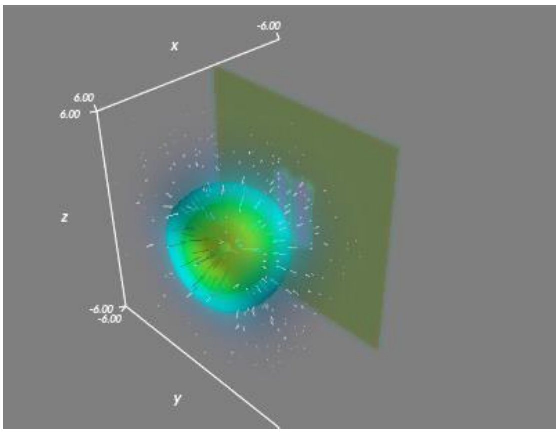
That is if the conjugate gradient method starts with $x_0 = 0$, then

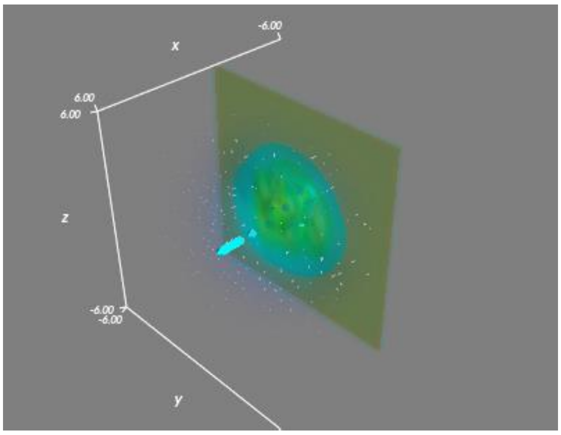
$$x_k = argmin_{y \in \mathbb{R}^N} \{(x-y)^T A(x-y) : y \in span\{b, Ab, \dots A^{k-1}b\}\}$$

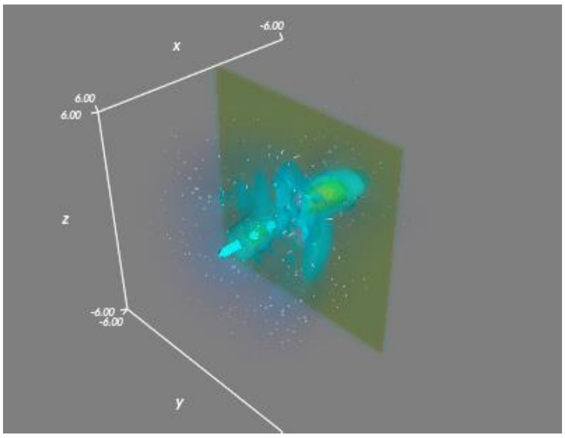
The algorithm below is detailed for solving $Ax = b$ where A is real, symmetric, positive definite matrix. The input vector x_0 can be an approximate initial solution or 0. It is a different formulation of the exact procedure described above.

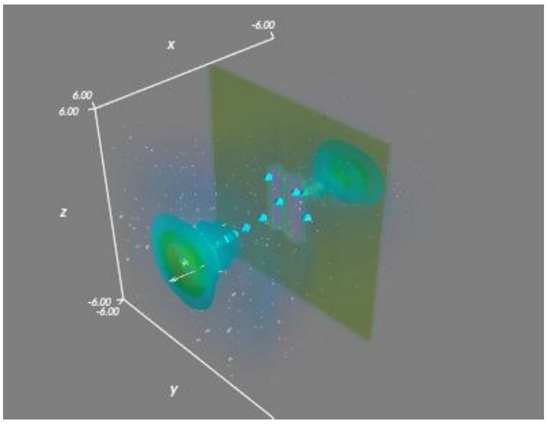
```
$r_0 := b - Ax_0$
If $r_0$ is sufficiently small, then return $x_0$ as the result, else
$p_0 := r_0$
$k := 0$
$repeat$
    $a_k := \frac{r_k^T r_k}{p_k^T A p_k}$
    $x_{k+1} := x_k + a_k p_k$
    $r_{k+1} := r_k - a_k A p_k$
    If $r_{k+1}$ is sufficiently small, then exit loop, else
    $\beta_k := \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$
    $p_{k+1} := r_{k+1} + \beta_k p_k$
    $k := k + 1$
end repeat
return $x_{k+1}$ as the result
```

The conjugate gradient method with a trivial modification is extendable to solving, given complex-valued matrix A and vector b, the system of linear equations $Ax = b$ for the complex-valued vector x, where A is Hermitian (i.e., A' = A) and positive-definite matrix, and the symbol ' denotes the conjugate transpose. The trivial modification is simply substituting the conjugate transpose for the real transpose everywhere.









Appendix.

ii. Series

The exponential function is given as

$$e^{x} = \sum_{n=0}^{\infty} \frac{x^{n}}{n!} = 1 + x + \frac{x^{2}}{2!} + \frac{x^{3}}{3!} + \cdots$$

And sine as

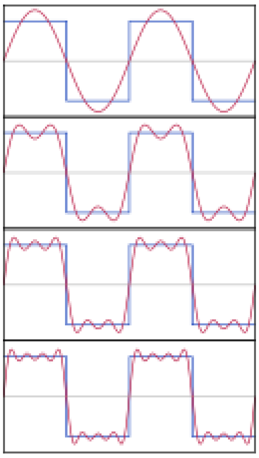
$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^{n} x^{2n+1}}{(2n+1)!} = x - \frac{x^{3}}{3!} + \frac{x^{5}}{5!} - \frac{x^{7}}{7!} + \cdots$$

These are all power series and are also all examples of Taylor Series.

A Fourier series is a continuous, periodic function created by a summation of harmonically related sinusoidal functions.

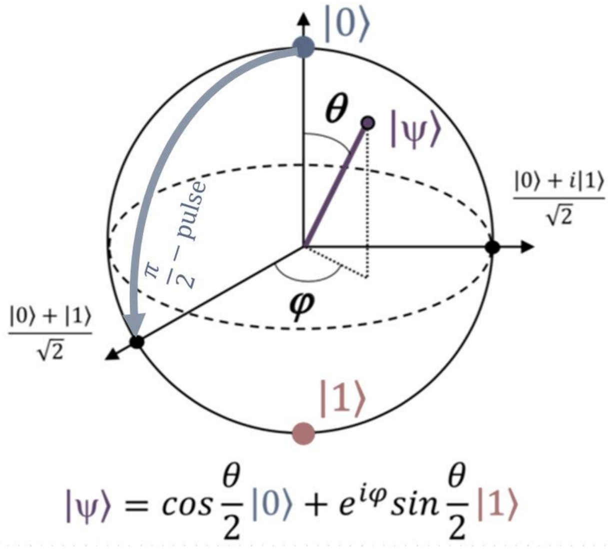
$$s_{n}(x) = A_{0} + \sum_{n=1}^{N} \left(A_{n} \cos\left(2\pi \frac{n}{P} x\right) + B_{n} \sin\left(2\pi \frac{n}{P} x\right) \right) = \sum_{n=-N}^{N} C_{n} e^{i 2\pi \frac{n}{P} x}$$

The first four partial sums of the Fourier series for a square wave. As more harmonics are added, the partial sums converge to (become more and more like) the square wave:



The Bloch or Poincare sphere representation for SU(2) where the base manifold is the S^2 sphere while the fibre is given by the U(1) phase at each point on that sphere. Together we have the finer bundle $SU(2) \cong S^{2} \times U(1)$. Thus, S^3 the 3-sphere, is a principal U(1)-bundle over the 2-sphere

The Bloch sphere is a geometrical representation of the pure state space $|\psi\rangle$ of a two-level quantum mechanical system. The Bloch sphere is a unit 2-sphere, with antipodal points corresponding to a pair of mutually orthogonal state vectors. The north and south poles of the Bloch sphere are typically chosen to correspond to the standard basis vectors $|0\rangle$ and $|1\rangle$, which in turn might correspond e.g. to the spin-up and spin-down states of an electron. The choice however is arbitrary. Given an orthonormal basis, any pure state $|\psi\rangle$ of a two-level quantum system can be written as a superposition of the basis vectors $|0\rangle$ and $|1\rangle$, where the



$|0\rangle$
θ
$|\psi\rangle$
$\frac{|0\rangle + i|1\rangle}{\sqrt{2}}$
$\frac{\pi}{2} - pulse$
φ
$\frac{|0\rangle + |1\rangle}{\sqrt{2}}$
$|1\rangle$
$|\psi\rangle = cos\frac{\theta}{2}|0\rangle + e^{i\varphi}sin\frac{\theta}{2}|1\rangle$

coefficient of (or contribution from) each of the two basis vectors is a complex number. This means that the state is described by four real numbers.

$\langle\psi|\psi\rangle = 1$, or equivalently $\parallel |\psi\rangle \parallel^2 = 1$, Given this constraint, that the pure state is unitary, we can write $|\psi\rangle$ using the following representation,

$|\psi\rangle = cos(\theta/2)|0\rangle + e^{i\phi}sin(\theta/2)|1\rangle = cos(\theta/2)|0\rangle + (cos\phi + isin\phi)sin(\theta/2)|1\rangle, where 0 \leq \theta \leq \pi and 0 \leq \phi < 2\pi$

The representation is always unique, because even though the value of ϕ is not unique when $|\psi\rangle$ is one of the states $|0\rangle$ or $|1\rangle$, the point represented by $\theta and \phi$ is unique.

i. Light at a point in Space as a complex number with real electric component and imaginary magnetic component. (The Silberstein Vector)

Heinrich Martin Weber published the fourth edition of "the partial differential equations of mathematical physics according to Riemann's lectures" in two volumes in 1900 and 1901. Weber demonstrated how to consolidate Maxwell's equations without charges or currents,

$$curl(E + iM) = \frac{i}{c}\frac{\partial(E + iM)}{\partial t}$$

This would be independently rediscovered in 1907 by Ludwik Silberstein.

Given an electric field E and a magnetic field B defined on a common region of spacetime, the Riemann-Silberstein vector is

$$F = E + icB$$

Where c is the speed of light ($c = 3 \cdot 10^8 ms^{-1}$). It is isomorphic to the electromagnetic tensor F, a 2-vector used in the covariant formulation of classical electromagnetism. The Riemann-Silberstein vector is used as a point of reference in the geometric algebra formulation of electromagnetism.

In a 1996 contribution to quantum electrodynamics Iwo Bialynicki-Birula used the Riemann-Silberstein vector as the basis for an approach to the photon, noting that it is a "complex vector-function of space coordinates and time that adequately describes the quantum state of a single photon." To put the Riemann-Silberstein vector in contemporary parlance, a transition is made, "With the advent of spinor calculus that superseded the

quaternionic calculus, the transformation properties of the Riemann-Silberstein vector have become even more transparent ... a symmetric second-rank spinor."
Maxwell's four equations in vector calculus reduce to one equation in the algebra of physical space.

$$(\frac{1}{c}\frac{\partial}{\partial t} + \nabla)F = \frac{1}{\epsilon_0}(\rho - \frac{1}{c}J)$$

The Riemann-Silberstein vector is a complex unit vector, $\hat{F} \in \mathbb{C}$, it is thus an element of the Clifford algebra, an algebra generated by a vector space with a quadratic form, that is a polynomial with terms all of degree 2 and is. a unitary associative algebra, $Cl_{0,1}(\mathbb{R})$,a two dimensional algebra generated by a single basis vector that squares to -1, and is algebra-isomorphic to C $(C \cong Cl_0(\mathbb{C}\})$. In contrast to the electron wave function the modulus square of the wave function of the photon (Riemann-Silberstein vector) is not dimensionless and must be multiplied by the "local photon wavelength" with the proper power to give dimensionless expression to normalise.

ii. *The sphere*

S^2, the 2-sphere, commonly the sphere, is the boundary of a 3-ball in 3-dimensional space.
A point on the two sphere can be parameterised with 3 cartesian coordinates $\sqrt{x^2 + y^2 + z^2} = 1$
Or by two angles, θ, ϕ, an angle in the x-y plane and one between the x-z, the angles of longitude and latitude.
$\hat{a} = (sin\theta cos\phi, sin\theta sin\phi, cos\theta) = (x, y, z)$

It is known that among the n-spheres, only $S^0, S^1 and\, S^3$ can admit a lie group structure, which is a highly non-trivial theorem.

$SU(2) = \{\begin{pmatrix}\alpha & -\overline{\beta}\\ \beta & \overline{\alpha}\end{pmatrix}: a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1\}$ Is diffeomorphic to the 3-sphere S^3, if we consider α, β as a pair in $\mathbb{C}^2$ where $\alpha = a + bi$ and $\beta = c + di$ then the equation $|a|^2 + |b|^2 = 1$ becomes $a^2 + b^2 + c^2 + d^2 = 1$. SU(2) is isomorphic group of versors. Quaternions of norm 1 are called versors since they generate the rotation group SO(3)

The SU(2) matrix

$$\begin{pmatrix}a + bi & c + di\\ -c + di & a - bi\end{pmatrix} \qquad (a, b, c, d \in \mathbb{R})$$

can be mapped to the quaternion $a\hat{1} + b\hat{i} + c\hat{j} + d\hat{k}$. This map is in fact the group isomorphism that relates SU(2) to versors.

One basis for SU(2) is given by, $t_1 = \frac{1}{2}\begin{bmatrix}0 & -i\\ -i & 0\end{bmatrix}, t_2 = \frac{1}{2}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}, t_3 = \frac{1}{2}\begin{bmatrix}-i & 0\\ 0 & i\end{bmatrix}$, These are related to the Pauli matrices by, $t_i \leftrightarrow \frac{1}{2i}\sigma_i$

The unitary group U(n) is a real Lie group of dimension n2. The Lie algebra of U(n) consists of n × n skew-Hermitian matrices.

In linear algebra it is often important to know which vectors have their directions unchanged by a linear transformation. An eigenvector or characteristic vector is such a vector. Thus an eigenvector v of a linear transformation T is scaled by a constant factor λ when the linear transformation is applied to it $Tv = \lambda v$.

Geometrically vectors are multi-dimensional quantities with magnitude and direction, often pictured as "arrows". A linear transformation rotates, stretches, or shears the vectors it acts upon. Its eigenvectors are those vectors that are only stretched with no rotation or shear. The corresponding eigenvalue is the factor by which an eigenvector is stretched or squished.

Hermitian matrices are fundamental to quantum mechanics because they describe operators with necessarily real eigenvalues. An eigenvalue a of an operator $\hat{A}$ on some quantum state $|\psi\rangle$ is one of the possible measurement outcomes of the operator, which requires the operators to have real eigenvalues.

A Hermitian matrix or self-adjoint complex square matrix that is equal to its own conjugate transpose— that is, the element in the ith row and and jth column is equal to the complex conjugate of the element in the jth row and ith column for all indices i and j.

$A_{Hermitian} \leftrightarrow a_{ij} = \overline{a_{ji}}$ or in matrix form $A_{Hermitian} \leftrightarrow A = \overline{A^T}$

The group U(2) is the set of all 2×2 matrices such that $U^\dagger U = I$.

Every matrix of the lie algebra has the form

$$aI + b_1\sigma_1 + b_2\sigma_2 + b_3\sigma_3$$

Where I is the 2×2 identity matrix, $\sigma_1, \sigma_2, \sigma_3$ are the three Pauli matrices, and $a_1, b_1, b_2, b_3 \in \mathbb{R}$.

This can be rewritten in the form

$$aI + b \cdot \sigma$$

Where $b = (b_1, b_2, b_3) \in \mathbb{R}^3$ and $\sigma = (\sigma_1, \sigma_2, \sigma_3)$

$$U(2) = \{\exp(iaI + ib \cdot \sigma) | a \in \mathbb{R} \; and \; b \in \mathbb{R}^3\}$$

Pauli matrices are Hermitian, involutory and unitary. Each Pauli matrix is Hermitian and together with the identity matrix I form a basis of the real vector space of 2 × 2 Hermitian matrices. This means that any 2 × 2 Hermitian matrix can be written in a unique way as a linear combination of Pauli matrices, with all coefficients being real numbers. Pauli matrices span the space of observables of the complex 2 dimensional Hilbert space. The Pauli matrices after multiplication by i to make them anti-hermitian also generate transformation in the sense of lie algebras: the matrices $i\sigma_1, i\sigma_2, i\sigma_3$ for a basis for the real lie algebra $SU(2)$, which exponentiates to the special unitary group SU(2). The algebra generated by the three matrices $\sigma_1, \sigma_2, \sigma_3$ are isomorphic to the Clifford algebra of $\mathbb{R}^3$, and the unitary associative algebra generated by $i\sigma_1, i\sigma_2, i\sigma_3$ are isomorphic to that of quaternions.

All three of the Pauli matrices can be compacted into a single expression:

$$\sigma_j = \begin{pmatrix} \delta_{j3} & \delta_{j1} - i\delta_{j_2} \\ \delta_{ji} + i\delta_{j2} & -\delta_{j3} \end{pmatrix}$$

δ_{jk} is the kronecker delta which equals +1 if j = k and 0 otherwise.

The matrices are involutory:

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = I_2$$

With the inclusion of the identity matrix, the Pauli matrices form an orthogonal basis of the Hilbert space of 2 × 2 Hermitian matrices, $\mathcal{H}_2$ over $\mathbb{R}$, and the Hilbert space of all complex 2 × 2 matrices over $\mathbb{C}$.

$$W^+ and \; W^- \; bosons \begin{cases} W_1 \\ W_2 \end{cases}$$
$$Z - boson \; and \; \gamma \quad \begin{cases} W_3 \\ B \end{cases}$$

Mathematically, electromagnetism is unified with the weak interactions as a Yang-Mills field with an $SU(2) \times U(1)$ gauge group, , which describes the formal operations that can be applied to the electroweak gauge fields without changing the dynamics of the system. These fields are the weak isospin fields W1, W2, and W3, and the weak hypercharge field B. This invariance is known as electroweak symmetry.

The generators of SU(2) and U(1) are given the name weak isospin T and weak hypercharge Y respectively. These then give rise to the gauge bosons that mediate the electroweak interactions – the three W bosons of weak isospin $(W_1, W_2, and \; W_3)$, and the B boson of weak hypercharge, respectively, all of which are "initially" massless. These are not physical fields yet, before spontaneous symmetry breaking and the associated Higgs mechanism.

In the Standard Model, the observed physical particles, the $W^{\pm}$ and Z_0 bosons, and the photon, are produced through the spontaneous symmetry breaking of the electroweak symmetry $SU(2) \times U(1)_Y$ to $U(1)_{em}$, effected by the Higgs mechanism (see also Higgs boson), an elaborate quantum-field-theoretic phenomenon that "spontaneously" alters the realization of the symmetry and rearranges degrees of freedom.

The electric charge arises as the particular linear combination (nontrivial) of YW (weak hypercharge) and the T3 component of weak isospin ($Q = T_3 + \frac{1}{2}Y_W$)that does not couple to the Higgs boson. That is to say: the Higgs and the electromagnetic field have no effect on each other, at the level of the fundamental forces ("tree level"), while any other combination of the hypercharge and the weak isospin must interact with the Higgs. This causes an apparent separation between the weak force, which interacts with the Higgs, and electromagnetism, which does not. Mathematically, the electric charge is a specific combination of the hypercharge and T3.

$U(1)_{em}$ (the symmetry group of electromagnetism only) is defined to be the group generated by this special linear combination, and the symmetry described by the $U(1)_{em}$ group is unbroken, since it does not directly interact with the Higgs.

The above spontaneous symmetry breaking makes the W3 and B bosons coalesce into two different physical bosons with different masses – the Z_0 boson, and the photon (γ),

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}$$

where θ_W is the weak mixing angle. The axes representing the particles have essentially just been rotated, in the (W_3, B) plane, by the angle θ_W. This also introduces a mismatch between the mass of the Z_0 and the mass of the $W^{\pm}$ particles (denoted as m_Z and m_W , respectively),

$$m_Z = \frac{m_W}{\cos\theta_W}$$

The W_1 and W_2 bosons, in turn, combine to produce the charged massive bosons $W^{\pm}$:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp iW_2)$$

Because the value of the mixing angle is currently determined empirically, in the absence of any superseding theoretical derivation it is mathematically defined as:

$$\cos\theta_w = \frac{m_W}{m_Z}$$

4.

The gamma matrices $\{\gamma^1, \gamma^2, \gamma^3, \gamma^4\}$ are a set of 4×4 complex matrices which satisfy the defining anti-commutation relations that ensure they generate a matrix representation of the Clifford algebra $Cl_{1,3}(\mathbb{R})$. When interpreted as the matrices of the action of a set of orthogonal basis vectors for contract variant vectors in Minkowski space, the column vectors on which the matrices act become a space of spinors, on which the Clifford algebra of spacetime acts, this in turn makes it possible to represent infinitesimal spatial rotations and Lorentz boosts.

γ^0 is the time-like hermitian matrix. The other three are space-like, anti-Hermitian matrices. More compactly $\gamma^0 = \sigma^3 \otimes I_2, and \gamma^j = i\sigma^2 \otimes \sigma^j$, where ⊗ denotes the Kronecker product (Block matrix multiplication maybe more detail needed).

The group SU(2) is the Lie group of unitary 2 × 2 matrices with unit determinant; its Lie algebra is the set of all 2 × 2 anti-Hermitian matrices with trace 0. The Lie algebra SU(2) is the 3-dimensional real algebra spanned by the set $i\sigma_k$

$SU(2) = span\{\frac{i\sigma_1}{2}, \frac{i\sigma_2}{2}, \frac{i\sigma_3}{2}\}$

The lie algebra SU(2) is isomorphic to the lie algebra SO(3), which corresponds to the Lie group SO(3), the group of rotations in three-dimensional space. In other words, one can say that the $i\sigma_j$ are a realisation and in fact, the lowest-dimensional realisation) of infinitesimal rotations in three-dimensional space. However even though SU(2) and SO(3) are isomorphic as Lie algebras, SU(2) and SO(3) are not isomorphic as Lie groups. SU(2) is actually a double cover of SO(3), meaning that there is a two-to-one group homomorphism from SU(2) to SO(3).

The real linear span of $\{I, i\sigma_1, i\sigma_2. i\sigma_3\}$ is isomorphic to the real algebra of quaternions, $\mathbb{H}$, represented by the span of the basis vectors $\{1, i, j, k\}$. The isomorphism from $\mathbb{H}$ to this set is given by the following map

$1 \mapsto I, i \mapsto -\sigma_2\sigma_3 = -i\sigma_1, j \mapsto -\sigma_3\sigma_1 = -i\sigma_2, k \mapsto -\sigma_1\sigma_2 = -i\sigma_3$

The spin group is used to describe the symmetries of (electrically neutral, unchanged) fermions. It's complexification, $Spin_c$ is used to describe electrically charged fermions, most notably the electron. Strictly speaking, the spin group describes a fermion in a zero-dimensional space; however, space is not zero-dimensional, and so the spin group is used to define spin structures on pseudo-Riemannian manifolds.

Each of the Hermitian Pauli matrices of spin$-\frac{1}{2}$ particles has two eigenvalues, +1 and -1. The corresponding normalised eigen vectors are,

$$\begin{aligned}
\psi_{x+} &= \left|\tfrac{1}{2}, \tfrac{+1}{2}\right\rangle_x = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, & \psi_{x-} &= \left|\tfrac{1}{2}, \tfrac{-1}{2}\right\rangle_x = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \\
\psi_{y+} &= \left|\tfrac{1}{2}, \tfrac{+1}{2}\right\rangle_y = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}, & \psi_{y-} &= \left|\tfrac{1}{2}, \tfrac{-1}{2}\right\rangle_y = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}, \\
\psi_{z+} &= \left|\tfrac{1}{2}, \tfrac{+1}{2}\right\rangle_z = \begin{pmatrix}1\\0\end{pmatrix}, & \psi_{z-} &= \left|\tfrac{1}{2}, \tfrac{-1}{2}\right\rangle_z = \begin{pmatrix}0\\1\end{pmatrix}.
\end{aligned}$$

By the postulates of quantum mechanics, an experiment designed to measure the electron spin on the x,y, or z axis can only yield an eigenvalue of the corresponding spin operator $(S_x, S_y, or S_z)$ on that axis, i.e. $\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$. The quantum state of a particle with respect to spin can be represented by a two-component spinor:

$\psi = \begin{pmatrix}a + bi\\c + di\end{pmatrix}$

… lots to talk about spinor

The Dirac equation is written in terms of a Dirac spinor field ψ taking values in a complex vector space described concretely as $\mathbb{C}^4$, defined on a flat space time $\mathbb{R}^{1,3}$.

In terms of a field $\psi: R^{1,3} \rightarrow \mathbb{C}^3$, the Dirac equation is then,

$$(i\hbar\gamma^{\mu}\partial_{\mu} - mc)\psi(x) = 0$$

$\{\gamma^u, \gamma^v\} = 2\eta^{uv}I_4$

These matrices are explicitly realised under a choice of representation, Pauli matrices are often chosen.

$\gamma^i = \begin{bmatrix}0 & \sigma^i\\-\sigma^i & 0\end{bmatrix}$

Square-integrable function L^2

$$f : \mathbb{R} \to \mathbb{C}\, square\ integrable \iff \int_{-\infty}^{\infty} |f(x)|^2 d\, x < \infty$$

Or more generally, as defined by the inner product,

$$\langle f, g\rangle = \int_A \overline{f(x)}\, g(x) dx$$

where,

f and g are square integrable functions,
$\overline{f(x)}$ is the complex conjugate of $f(x)$,
since $|a|^2 = a \cdot \bar{a}$, square integrability is the same as $\langle f, f\rangle < \infty$

An equivalent definition is to say that the square of the function itself (rather than of its absolute value) is Lebesgue integrable. For this to be true, the integrals of the positive and negative portions of the real part must both be finite, as well as those for the imaginary part.

The space of square integrable functions is the L^p space in which $p = 2$.

The vector space $\mathbb{R}^n$, together with the dot product, is an example of an inner product space. In particular, $\mathbb{R}$ itself (as a vector space), with multiplication, is an inner-product space.